

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 08:14:32 ; Search time 20 Seconds  
(without alignments)  
817.628 Million cell updates/sec

Title: US-09-171-607A-1

Perfect score: 893

Sequence: 1 VALNSPLSGGMRGIRGADPQ.....ASCHHAYIVLCIENSFMTAS 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_78:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	893	100.0	684	2 A53019	collagen alpha 1(X
2	778	87.1	1774	2 B56101	collagen alpha 1(X
3	775	86.8	1315	2 A56101	collagen alpha 1(X
4	504	56.4	1388	2 A53317	collagen alpha 1(X
5	357	40.0	650	2 T22002	hypothetical prote
6	79	8.8	477	2 T05202	pectinesterase hom
7	77	8.6	314	2 AG3054	succinoglycan bios
8	77	8.6	314	2 F98231	succinoglycan bios
9	76	8.5	204	2 B45022	CRK-I - human
10	76	8.5	636	2 J00047	class I cytochrome
11	75.5	8.5	904	2 T04377	probable pullulana
12	75	8.4	351	2 B89781	conserved hypothet
13	74.5	8.3	850	2 S56015	gastric mucin MUC5
14	73.5	8.2	1653	2 B91052	hypothetical prote
15	73.5	8.2	1653	2 B9896	hypothetical prote
16	73.5	8.2	1653	2 J05028	hypothetical prote
17	73	8.2	435	2 J05028	hyaluronoglycosam
18	73	8.2	462	2 T17480	endo-xylanase homo
19	73	8.2	513	1 A45333	exopolysphatase
20	73	8.2	513	2 D91049	exopolysphatase
21	73	8.2	513	2 H85893	exopolysphatase
22	72.5	8.1	884	2 C70729	hypothetical prote
23	72	8.1	239	2 A46243	epidermal growth f
24	72	8.1	297	2 AE3128	carbohydrate kinas
25	72	8.1	304	2 A45022	CRK-II - human
26	72	8.1	304	2 I58394	c-Crk - mouse
27	72	8.1	319	2 D98159	hypothetical prote
28	72	8.1	800	2 AG0438	probable Rhs acces
29	72	8.1	1006	2 T00050	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

A53019  
collagen alpha 1(XVIII) chain - human (fragment)  
N:Contains: endostatin  
C:Species: Homo sapiens (man)  
C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 15-Sep-2003  
C:Accession: A53019  
R:Oh, S.P.; Warman, M.L.; Seldin, M.F.; Cheng, S.D.; Knoll, J.H.M.; Timmons, S.; Olsen  
Genomics 19, 494-499, 1994  
A:Title: Cloning of cDNA and genomic DNA encoding human type XVIII collagen and locali:  
A:Reference number: A53019; MUID:94245237; PMID:8188291  
A:Accession: A53019  
A:Molecule type: mRNA  
A:Residues: 1-684 <OHA>  
A:Cross-references: GB:L22548; NID:G348908; PIDN:AAA51864.1; PID:G562794  
A:Note: The cited accession number, L25548, is not in Genbank release 103  
A:Note: in the authors' translation, 482-gly is not shown, residues 483-490 are shifte  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
lated and subsequently O-glycosylated.  
C:Comment: Different splice forms of collagen alpha 1(XVIII) may be involved in periva:  
C:Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of  
ay be useful in treating solid tumors.

##### C:Genetics:

A:Gene: GDB:COL18A1  
A:Cross-references: GDB:138752; OMIM:120328  
A:Map position: 21q22.3-21q22.3  
C:Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteogl;  
F:1-684/Product: collagen alpha 1(XVIII) chain (fragment) #status predicted <MAT>  
F:1-59/Domain: collagenous (fragment) #status predicted <CO4>  
F:74-115/Domain: collagenous #status predicted <CO5>  
F:123-201/Domain: collagenous #status predicted <CO6>  
F:212-244/Domain: collagenous #status predicted <CO7>  
F:257-278/Domain: collagenous #status predicted <CO8>  
F:262-264/Region: cell attachment (R-G-D) motif  
F:286-340/Domain: collagenous #status predicted <CO9>  
F:354-371/Domain: collagenous #status predicted <CO10>  
F:502-684/Product: endostatin #status predicted <EST>  
F:509-684/Region: multiplexin collagen carboxyl-terminal homologous

Query Match 100.0%; Score 893; DB 2; Length 684;

Best Local Similarity 100.0%; Pred. No. 1e-81;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGGMRGIRGADPQFCQARAVGLAGTFRFLSSRLQDLIYSIVRRADRAAIVP 60

DB 514 VALNSPLSGGMRGIRGADPQFCQARAVGLAGTFRFLSSRLQDLIYSIVRRADRAAIVP 573

QY 61 NLKDELLFPSPSEALFSGSEGLKPGARIFSGDKDLRHPTWPKQSVWHGSDPNRRRLTE 120

DB 574 NLKDELLFPSPSEALFSGSEGLKPGARIFSGDKDLRHPTWPKQSVWHGSDPNRRRLTE 633

QY 121 SYCETWRTAPSATGQASLLGLGRLLGQSAASCHHAYIVLCIENSFMTAS 170

```

|||||
634 SYCETWTEAPATGQASSLLGRLGQAASCHHAYIVLCIENSFWTAS 683
|||||

Db
RESULT 2
B56101
collagen alpha 1(XVIII) chain precursor, long splice form - mouse
N:Contains: collagen alpha 1(XVIII) chain precursor, medium splice form; endostatin
C:Species: Mus musculus (house mouse)
C>Date: 03-Oct-1995 #sequence revision 08-May-1998 #text change 15-Sep-2003
C:Accession: B56101.1; C56101; S72450; S65595; PNO675; A54072; A58816
R:Rehn, M.; Pihlajaniemi, T.
J. Biol. Chem. 270, 4705-4711, 1995
A:Title: Identification of three N-terminal ends of type XVIII collagen chains and tis-
tif homologous to rat and Drosophila frizzled proteins.
A:Reference number: A56101; MUID:95181468; PMID:7876242
A:Accession: B56101
A:Molecule type: mRNA
A:Residues: 1-562 <REH1>
A:Cross-references: GB:U11637; NID:G618429; PIDN:AAC52179.1; PID:G618430
A:Experimental source: splice form clone PE17.24
A:Accession: C56101
A:Molecule type: mRNA
A:Residues: 1-239,487-562 <REH2>
A:Cross-references: GB:U11637; NID:G618429
A:Experimental source: splice form clones PB8.1, PE19, PE15.2
R:Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.
submitted to the EMBL Data Library, August 1993
A:Reference number: S72450
A:Accession: S72450
A:Molecule type: mRNA
A:Residues: 487-1146, 'L', 1148-1193, 'F', 1195-1210, 'R', 1212-1512, 'L', 1514-1522, 'F', 1524-16
A:Cross-references: EMBL:L22545; NID:G348968; PIDN:AAA19787.1; PID:G511298
R:Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.
Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994
A:Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa
A:Reference number: A58370; MUID:94240111; PMID:8183893
A:Accession: S65595
A:Molecule type: mRNA
A:Residues: 487-1512, 'L', 1514-1522, 'F', 1524-1683, 'V', 1685-1774 <OH2>
A:Cross-references: EMBL:L22545
R:Abe, N.; Muragaki, Y.; Yoshitaka, H.; Inoue, H.; Ninomiya, Y.
Biochem. Biophys. Res. Commun. 196, 576-582, 1993
A:Title: Identification of a novel collagen chain represented by extensive interruptions
A:Reference number: PNO675; MUID:94059075; PMID:8240330
A:Accession: PNO675
A:Molecule type: mRNA
A:Residues: 635-1774 <ABE>
R:Rehn, M.; Hintikka, E.; Pihlajaniemi, T.
J. Biol. Chem. 269, 13929-13935, 1994
A:Title: Primary structure of the alpha1 chain of mouse type XVIII collagen, partial str
collagen chain.
A:Reference number: A54072; MUID:94245707; PMID:8188673
A:Accession: A54072
A:Molecule type: DNA; mRNA
A:Residues: 1293-1403, 'R', 1405-1774 <REH3>
A:Cross-references: GB:U03714; NID:G487733; PIDN:AAA20657.1; PID:G487734
R:O'Reilly, M.S.; Boehm, T.; Shing, Y.; Fukai, N.; Vasios, G.; Lane, W.S.; Flynn, E.; B
Cell 88, 277-285, 1997
A:Title: Endostatin: an endogenous inhibitor of angiogenesis and tumor growth.
A:Reference number: A58816; MUID:97160848; PMID:9008168
A:Accession: A58816
A:Molecule type: protein
A:Residues: 1591-1610 <ORE>
A:Experimental source: hemangi endothelium cells
A:Note: inhibits endothelial cell proliferation
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
lated and subsequently O-glycosylated.
C:Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in peri
C:Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of un
ay be useful in treating solid tumors.
C:Genetics:
A:Gene: MGI:Coll18a1

```

A:Cross-references: MGI:71175  
A:Map position: 10:41.0  
A:Introns: 1295/3; 1310/1; 1331/1; 1345/3; 1388/3; 1437/1; 1461/3; 1505/3; 1516/3; 159;  
A:Note: the list of introns is incomplete  
C:Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteogly-  
F1-1774/Product: collagen alpha 1(XVIII) chain precursor, long splice form #status pri-  
F1-239,487-1774/Product: collagen alpha 1(XVIII) chain precursor, medium splice form  
F1-24/Domain: signal sequence #status predicted <SIG>  
F1786-812/Domain: collagenous #status predicted <CO01>  
F1823-896/Domain: collagenous #status predicted <CO02>  
F1921-1042/Domain: collagenous #status predicted <CO03>  
F1066-1148/Domain: collagenous #status predicted <CO04>  
F1163-1204/Domain: collagenous #status predicted <CO05>  
F1218-1290/Domain: collagenous #status predicted <CO06>  
F1301-1333/Domain: collagenous #status predicted <CO07>  
F1346-1369/Domain: collagenous #status predicted <CO08>  
F1351-1353/Region: cell attachment (R-G-D) motif  
F1377-1428/Domain: collagenous #status predicted <CO09>  
F1442-1459/Domain: collagenous #status predicted <CO10>  
F1591-1774/Product: endostatin #status predicted <EST>  
F1598-1774/Region: multiplexin collagen carboxyl-terminal homologous  
F1354,361,947/Binding site: carboxylate (Asn) (covalent) #status predicted  
F1699,704,1716/Binding site: carboxylate (Ser) (covalent) #status predicted  
F1910,913,1053/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 87.1%; Score 778; DB 2; Length 1774;  
Best Local Similarity 85.8%; Pred. No. 1.2e-69;  
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNPLSGMGRGIRGADPQCFQARAVGLATGTFRAFLSSRLQDLYSIYRRADRAVPV 60  
|||||  
Db 1603 VALNPLSGMGRGIRGADPQCFQARAVGLSTGTFRAFLSSRLQDLYSIYRRADRGVPIV 1662  
|||||

QY 61 NLKDELLPSPSWEALPSGSGPLKPGARIPSPGKQVLRHTWPKQSVHSGSPNGRRLTE 120  
|||||  
Db 1663 NLKDEVLSFWSLPSGSGQCPGARIPSPGKQVLRHTWPKQSVHSGSPNGRRLME 1722  
|||||

QY 121 SYCETWTEAPATGQASSLLGRLGQAASCHHAYIVLCIENSFWTAS 169  
|||||  
Db 1723 SYCETWTEATGATGQASSLLGRLLEKQAASCHNSYIVLCIENSFWTAS 1771  
|||||

RESULT 3  
A56101  
collagen alpha 1(XVIII) chain precursor, short splice form - mouse  
N:Contains: endostatin  
C:Species: Mus musculus (house mouse)  
C>Date: 03-Oct-1995 #sequence revision 08-May-1998 #text change 15-Sep-2003  
C:Accession: A56101; A58371; S72450; S65595  
R:Rehn, M.; Pihlajaniemi, T.  
J. Biol. Chem. 270, 4705-4711, 1995  
A:Title: Identification of three N-terminal ends of type XVIII collagen chains and tis-  
tif homologous to rat and Drosophila frizzled proteins.  
A:Reference number: A56101; MUID:95181468; PMID:7876242  
A:Accession: A56101  
A:Molecule type: mRNA  
A:Residues: 1-103 <REH1>  
A:Cross-references: GB:U11636; NID:G618427; PIDN:AAC52178.1; PID:G618428  
R:Rehn, M.; Pihlajaniemi, T.  
Proc. Natl. Acad. Sci. U.S.A. 91, 4234-4238, 1994  
A:Title: Alpha1(XVIII), a collagen chain with frequent interruptions in the collagenous  
A:Reference number: A58371; MUID:94240112; PMID:8183894  
A:Accession: A58371  
A:Molecule type: mRNA  
A:Residues: 1-928 <REH2>  
A:Cross-references: GB:L16898; NID:G404754; PIDN:AAA37434.1; PID:G553894  
R:Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.  
submitted to the EMBL Data Library, August 1993  
A:Reference number: S72450  
A:Accession: S72450  
A:Molecule type: mRNA  
A:Residues: 28-687, 'L', 689-734, 'F', 736-751, 'R', 753-1315 <OHW>  
A:Cross-references: EMBL:L22545; NID:G348968; PIDN:AAA19787.1; PID:G511298

R; Muragaki, Y.; Abe, N.; Ninomiya, Y.; Olsen, B.R.; Ooshima, A.  
J. Biol. Chem. 269, 4042-4046, 1994  
A; Title: The human alpha(XV) collagen chain contains a large amino-terminal non-trip  
A; Reference number: A53146; MUID: 94140817; PMID: 8307960  
A; Accession: A53146  
A; Status: preliminary  
A; Molecule type: mRNA  
A; Residues: 1-9, 'S', 11-48, 'V', 50-94, 'A', 96-149, 'A', 151-203, 'V', 205-408, 'A', 410-569 <MUI  
A; Cross-references: GB:D21230; NID:G415605; PID:BA04762.1; PID:d1005294; PID:g460703  
R; Myers, J.C.; Kivirikko, S.; Gordon, M.K.; Dion, A.S.; Pihlajaniemi, T.  
Proc. Natl. Acad. Sci. U.S.A. 89, 10144-10148, 1992  
A; Title: Identification of a previously unknown human collagen chain, alpha(XV), charac  
A; Reference number: S28778; MUID: 93066196; PMID: 1279671  
A; Accession: S28778  
A; Status: preliminary  
A; Molecule type: mRNA  
A; Residues: 544-640, 'P', 642-811, 'P', 813-1252 <MYE>  
C; Genetics:  
A; Gene: GDB:COL15A1  
A; Cross-references: GDB:l132578; OMIM:120325  
A; Map position: 9q21-9q22  
F; 1-22/Domain: signal sequence #status predicted <SIG>  
F; 23-1388/Product: collagen alpha 1(XV) chain #status predicted <MAT>  
F; 1216-1388/Region: multiplexin collagen carboxyl-terminal homologous

Query Match 56.4%; Score 504; DB 2; Length 1388;  
Best Local Similarity 56.9%; Pred. No. 3.3e-42;  
Matches 95; Conservative 27; Mismatches 41; Indels 4; Gaps 1;  
Db 1222 ALNPFSGDIR---ADPQCFKQARAAGLLSTYRAFLSSHLQDLSTIVRKAERYSLPIYN 1277

Qy 2 ALNSPLSGMGRIGADPQCQPARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIYN 61  
Db 1222 ALNPFSGDIR---ADPQCFKQARAAGLLSTYRAFLSSHLQDLSTIVRKAERYSLPIYN 1277

Qy 62 LKDELLPSPWALPGSGEGPLKPGARISEFOCKQVLRHPTWPQKSVWHGSDPNGRRLTES 121  
Db 1278 LKGQVLFNNWDSIYSGHGQGFNMHIPIYSEFGDRDITDTPSWFQKVIWHGSSPHGVRLVDN 1337

Qy 122 YCETWRTEAPSATGQASILGRLGQGAASCHHAYIVLICIENSFMT 168  
Db 1338 YCEAWRTADTAVTGLASPLSTGKILDQKAYSCANRLIVLICIENSFMT 1384

RESULT 5  
T22002  
hypothetical protein F39H11.4 - Caenorhabditis elegans  
C; Species: Caenorhabditis elegans  
C; Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text\_change 15-Oct-1999  
C; Accession: T22002  
R; White, S.  
submitted to the EMBL Data Library, October 1996  
A; Reference number: Z19500  
A; Accession: T22002  
A; Status: preliminary; translated from GB/EMBL/DDBJ  
A; Molecule type: DNA  
A; Residues: 1-650 <WIL>  
A; Cross-references: EMBL:Z81079; PIN: CAB03084.1; GSPDB: GN00019; CESP: F39H11.4  
A; Experimental source: clone F39H11  
C; Genetics:  
A; Gene: CESP:F39H11.4  
A; Map position: 1  
A; Introns: 109/3; 154/1; 357/1; 420/3; 464/3; 566/2; 594/1; 628/3

Query Match 40.0%; Score 357; DB 2; Length 650;  
Best Local Similarity 44.0%; Pred. No. 8.1e-28;  
Matches 73; Conservative 24; Mismatches 63; Indels 6; Gaps 4;  
Qy 1 VALNSPLSGMGRIGADPQCQPARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 60  
Db 468 IALSOPFSGNLHGLRGADLQCYREARAAGYTTTFRAMLSNVQDLVRIHVSVD-FDTTVV 526

Qy 61 NLKDELLFPSPWALPGSGEGPLKPGARISEFGCKVLRHPTWPQKSVWHGSDPNGRRLTE 120  
502 NTKCHLTFSPGSEFVNGAQQ--NNPDAVTFSEDFPDVINDSEWPKRVWHGSKDGGIR-AE 583



```

RESULT 10
JW0047
class I cytokinase receptor precursor - human
N:Alternate names: WSX-1
C:Species: Homo sapiens (man)
C:Date: 17-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000
C:Accession: JW0047
R:Spacher, C.A.; Grant, F.J.; Baumgartner, J.W.; Presnell, S.R.; Schrader, S.K.; Yamagi
Biochem. Biophys. Res. Commun. 246, 82-90, 1998
A:Title: Cloning and characterization of a novel class I cytokine receptor.
A:Reference number: JW0047; PMID:98262921; PMID:9600072
A:Accession: JW0047
A:Molecule type: mRNA
A:Residues: 1-636 <SPR>
A:Cross-references: GB:AF053004; NID:G3153240; PIDN:AAC39755.1; PID:G3153241
C:Genetics:
A:Map position: 19p13.11
C:Keywords: glycoprotein
F:1-32/Domain: signal sequence #status predicted <SIG>
F:515-540/Domain: transmembrane #status predicted <TM>
F:554-561/Domain: cytoplasmic #status predicted <CTP>
F:51,76,302,311,374,382,467/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 8.5%; Score 76; DB 2; Length 636;
Best Local Similarity 24.2%; Pred. No. 15;
Matches 39; Conservative 12; Mismatches 46; Indels 64; Gaps 7;
QY 11 MRGIRGADF-----OCFQARAVGLAGTFRAPLSSRLQDL----- 45
Db 1 MRGGRGAPFWLPLKALLLPILWLIVFQTRPQSGAGPLQCYGVGLGDLNCSWEPLGLDL 60
QY 46 -----YSIVRRADRAAVPT-----VNLKDELLF-----PSWEALFS 76
Db 61 GAPSELHLQSKYRSNKTQTVAAGRSWVAIPRELQIWSKLLVWGTACAGPLPPVVFV 120
QY 77 GSEGLPKFGA-RI-----FSGDKGVLR-----HPTWPKQSV 107
Db 121 NLETKMKNAPRLGPDVDFSEDDPLEATVHWAPTPWPSHKV 161
RESULT 11
T04377
probable pullulanase (EC 3.2.1.41) - barley
N:Alternate names: pullulanase
C:Species: Hordeum vulgare (barley)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Jun-2003
C:Accession: T04377
R:Lok, F.; Kristensen, M.; Planchot, V.; Leah, R.; Svendsen, I.; Svendsen, B.
submitted to the EMBL Data Library, December 1997
A:Description: Isolation and characterization of starch debranching enzyme, limit dextrin
A:Reference number: Z15320
A:Accession: T04377
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-904 <LOK>
A:Cross-references: EMBL:AF022725; NID:G2502057; PIDN:AAD04189.1; PID:G2677837
A:Experimental source: cv. Igri
C:Genetics:
A:Gene: HvLDP99
A:Introns: 21/3; 70/1; 87/3; 124/1; 169/3; 235/2; 285/2; 305/3; 346/1; 376/2; 407/3; 439
C:Superfamily: pullulanase type debranching enzyme
C:Keywords: glycosidase; hydrolase
Query Match 8.5%; Score 75.5; DB 2; Length 904;
Best Local Similarity 22.5%; Pred. No. 25;
Matches 45; Conservative 19; Mismatches 67; Indels 69; Gaps 8;
QY 5 SPLSGMRGIRGADFQCFQARAVGLAGT-----FPAFLSSRLQDLYSIVR----- 50
Db 62 SPSNG---GIQGYDSKVELQPSAGLPTVTQKPFIFSSYRAFKVPSVDVASLVKQLV 118

```

```

QY 51 -----RADRAAVPIVNL-----KDELLFPSWE-----ALFSG 77
Db 119 VASFGADGKHVDVTGTLQPLGVLDMDFAVTGPLGAVFSEDSVSLHLMWAPTAQGVSVCFDQ 178
QY 78 SEGP-----LKPGARIFSPGK-----DVLRHPTWPKQSVMHGSDPPNGRRIT 119
Db 179 PAGPALETVOLKESNGVSVTGPREWENRYLYEVDVY-HPTKAQVLKCLAGDPYARSLS 237
QY 120 ESYCETWTEAPSATQOASS 139
Db 238 ANGARTWLVDINNETLKPAS 257
RESULT 12
B89781
conserved hypothetical protein SA0184 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89781
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ono,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: B89781
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-351 <KUR>
A:Cross-references: GB:BA000018; PID:G13700106; PIDN:BA841405.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA0184
Query Match 8.4%; Score 75; DB 2; Length 351;
Best Local Similarity 23.3%; Pred. No. 9.2;
Matches 42; Conservative 17; Mismatches 67; Indels 54; Gaps 8;
QY 11 MRGIRGADFQCFQARAVGLAGTFRAPLSSRLQDL-----YSIVRRADRA-AVPTV 60
Db 97 IEATMAQGLKCLNASTIS-----RELTSLHQQLNDFTLSPFCHNYYPRTDTGLSVDLV 151
QY 61 NLKDELLFPSWEALFSGSEGLPKPGARIFSPGDKDVLRLHPTWPKQSVMHGSDPPNGRRLTE 120
Db 152 MKKNELIV-----QFNPKAQIYGVIGVSGLRGPL-----HKGLPT----- 186
QY 121 SYCETWTEAP-----SATQOASSLLGRLGQSAAS-----CHHAYIVLCIENSFMT 168
Db 187 --IEATRHSHPVAAKLLQETGVSEVLVGSLSIEMRQAKQLIDFCRHRHFTLCIEEVDPT 244
RESULT 13
S56015
gastric mucin MUCSAC - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Apr-2000
C:Accession: S56015; S53361
R:Klomp, L.W.J.; van Rens, L.; Strous, G.J.
Biochem. J. 308, 831-838, 1995
A:Title: Cloning and analysis of human gastric mucin cDNA reveals two types of conserve
A:Reference number: S56015; PMID:97104281; PMID:8948439
A:Accession: S56015
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-850 <KLO>
A:Cross-references: EMBL:X81649; NID:G547516; PIDN:CAA57309.1; PID:G547517
R:Guyonnet-Duperat, V.; Audie, J.P.; Debailleul, V.; Laine, A.; Buisine, M.P.; Galieque
Biochem. J. 305, 211-219, 1995
A:Title: Characterization of the human mucin gene MUCSAC: a consensus cysteine-rich dom
A:Reference number: S53361; PMID:95126907; PMID:7826332
A:Accession: S53361
A:Status: preliminary; nucleic acid sequence not shown

```

A:Molecule type: mRNA  
A:Residues: 648-678, 'L', 680-733, 'L', 735-760 <GUY>  
A:Cross-references: EMBL:Z34280; NID:G563380; PIDN:CAA84034.1; PID:G563381  
A:Experimental source: clone JUL32  
A>Note: this publication is not cited in GenBank entry HSMUCIN5, release 113.0

Query Match 8.3%; Score 74.5; DB 2; Length 850;  
Best Local Similarity 24.4%; Pred. No. 30;  
Matches 39; Conservative 23; Mismatches 71; Indels 25; Gaps 6;

QY 13 GIRGADQFCQQAARAVGLAGTFRALSSRLQDLYSIVRRADR-AAVPIVNLKDELLFPSPW 71  
Db 592 GINGGDFDFQNRDEGY--TF-----CESPSVQCHAESFNPPLADLGDVLCISHT 642  
QY 72 EALFSGSEGLKP-----GARIFSPGKDVLRHPTWPQSVW-----HGSDPNGRRLTE 120  
Db 643 EGLICLNKQLPPICYNYEIRIQCCETVNVCRDITRPKTVATTRPHTGAQTQTTF 702  
QY 121 SYCETWRTAPSATGQ-----ASSLLGRLLGQSAASCH 154  
Db 703 THMPSASTEQPTATSRGGPTATSVTQCTHTPVRNCH 740

## RESULT 14

B91052  
hypothetical protein ECs3386 [imported] - Escherichia coli (strain O157:H7, substrain R1)  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C/Accession: B91052  
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yaenunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: B91052  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1653 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA836809.1; PID:G13362856; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain R1MD 0509952  
C:Genetics:  
A:Gene: ECs3386

Query Match 8.2%; Score 73.5; DB 2; Length 1653;  
Best Local Similarity 24.4%; Pred. No. 84;  
Matches 32; Conservative 21; Mismatches 47; Indels 31; Gaps 5;

QY 29 GLAGTFRALSSRLQDLY-----SIVRRADRAAVPIVNLKDELLFPSPWEALFSGSE 79  
Db 380 GAPGYSKQFFMFGPRDLRPGETVILNGLLRDADGKALPNQPIKLDVIKPDGQVLSVVS 439  
QY 80 GPLKPGARIFSPGKDVLRHPTWPQKS-----VWH---GSDPNGRRLTESYCTWRT-- 129  
Db 440 QP-----ENGLYHFTWPLDSNAATGCMWHIRANTGDNQYRMWDFHVEDFMPERM 487  
QY 130 APSATGQASSL 140  
Db 488 ALNLTGKTKPL 498

## RESULT 15

F85896  
hypothetical protein Z3787 [imported] - Escherichia coli (strain O157:H7, substrain EDL9  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C/Accession: F85896  
R/Ferna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: F85896  
A>Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-1653 <STO>  
A:Cross-references: GB:AB005174; NID:G12516921; PIDN:AAGS7634.1; GSPDB:GN00145; UWGP:Z:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z3787

Query Match 8.2%; Score 73.5; DB 2; Length 1653;  
Best Local Similarity 24.4%; Pred. No. 84;  
Matches 32; Conservative 21; Mismatches 47; Indels 31; Gaps 5;

QY 29 GLAGTFRALSSRLQDLY-----SIVRRADRAAVPIVNLKDELLFPSPWEALFSGSE 79  
Db 380 GAPGYSKQFFMFGPRDLRPGETVILNGLLRDADGKALPNQPIKLDVIKPDGQVLSVVS 439  
QY 80 GPLKPGARIFSPGKDVLRHPTWPQKS-----VWH---GSDPNGRRLTESYCTWRT-- 129  
Db 440 QP-----ENGLYHFTWPLDSNAATGCMWHIRANTGDNQYRMWDFHVEDFMPERM 487  
QY 130 APSATGQASSL 140  
Db 488 ALNLTGKTKPL 498

Search completed: March 13, 2004, 08:18:27  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 08:10:51 ; Search time 18 Seconds  
(without alignments)  
491.774 Million cell updates/sec

Title: US-09-171-607A-1

Perfect score: 893  
Sequence: 1 VALNSPLSGMGRGADPQ.....ASCHRAYIVLCIENSFWTAS 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	893	100.0	1516	1	CALH_HUMAN
2	778	87.1	1774	1	CALH_MOUSE
3	504	56.4	1388	1	CALF_HUMAN
4	76	8.5	556	1	GLI_CHICK
5	75.5	8.5	334	1	XA5L_MOUSE
6	75	8.4	271	1	Y4BG_RHISN
7	74	8.3	296	1	CRK_XENLA
8	74	8.3	780	1	STRN_RAT
9	73.5	8.2	325	1	XA5L_HUMAN
10	73.5	8.2	1653	1	YFHM_ECOLI
11	73	8.2	512	1	PPX_ECOLI
12	72.5	8.1	884	1	YP67_MYCTU
13	72.5	8.1	884	1	YP97_MYCBO
14	72	8.1	304	1	CRK_HUMAN
15	72	8.1	304	1	CRK_MOUSE
16	72	8.1	304	1	CRK_RAT
17	72	8.1	613	1	HS75_CANAL
18	71	8.0	787	1	OXAA_CHLMO
19	70	7.8	780	1	STRN_MOUSE
20	70	7.8	837	1	LSTR_MOUSE
21	70	7.8	953	1	CAR4_MOUSE
22	69	7.7	266	1	CB21_SINAI
23	69	7.7	309	1	YHCC_ECOLI
24	69	7.7	326	1	TMOE_PSEME
25	69	7.7	462	1	MYCN_MOUSE
26	69	7.7	494	1	ALG8_PSEAE
27	69	7.7	1233	1	NWE3_HUMAN
28	68.5	7.7	390	1	PKG_EUCAL
29	68.5	7.7	1289	1	CSAB_BACUD
30	68	7.6	610	1	MET7_SCHPO
31	68	7.6	953	1	CAR4_HUMAN
32	68	7.6	999	1	MERK_HUMAN
33	67.5	7.6	348	1	NU2M_BRARE

## RESULT 1

CALH\_HUMAN STANDARD; PRT; 1516 AA.

ID CALH\_HUMAN Q9Y6Q7; Q9Y6Q8;

AC P39060; Q9Y6Q7; Q9Y6Q8;

DT 01-FEB-1995 (Rel. 31, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Collagen alpha 1(XVII) chain precursor [Contains: Endostatin].

GN COL18A1

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98164036; PubMed=9503365;

RA Saarela J., Ylikarppa R., Rehn M., Purmonen S., Pihlajaniemi T.;

RT "Complete primary structure of two variant forms of human type XVIII

collagen and tissue-specific differences in the expression of the

corresponding transcripts.";

RL Matrix Biol. 16:319-328(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=20289799; PubMed=10830953;

RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,

RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,

RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,

RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,

RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,

RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,

RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,

RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,

RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,

RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,

RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,

RA Lehrach H., Reinhardt R., Yaspo M.-L.;

RT "The DNA sequence of human chromosome 21.";

RL Nature 405:311-319(2000).

RN [3]

RP SEQUENCE OF 834-1516 FROM N.A.

RX MEDLINE=94245237; PubMed=8188291;

RA Oh S.P., Warman M.D., Seidman M.F., Cheng S., Knoll J.H., Timmons S.,

RA Olsen B.R.;

RT "Cloning of cDNA and genomic DNA encoding human type XVIII collagen

and localization of the alpha 1(XVII) collagen gene to mouse

chromosome 10 and human chromosome 21.";

RL Genomics 19:494-499(1994).

RN [4]

RP SEQUENCE OF 1334-1516 FROM N.A.

RX TISSUE=Placenta;

RA Zhi-Yong H., Biao L., Wei-Jie Z., Xiang-Fu W.;

RT "Cloning and expression of human endostatin gene in Escherichia

coli.";

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RN [5]

RP INVOLVEMENT IN KNOBLOCH SYNDROME.

Q9h257 homo sapien  
O08914 mus musculus  
P94281 bartonella  
P18909 rana catesb  
P17482 homo sapien  
Q10512 mycobacteri  
Q9jiv1 mus musculus  
Q8k1m5 rattus norv  
O60755 homo sapien  
P24295 clostridium  
Q62240 mus musculus  
Q72407 homo sapien

## ALIGNMENTS



RX MEDLINE=20400145; PubMed=10942434;  
 RA Sertie A.L., Sossi V., Camargo A.A., Zatz M., Brahe C.,  
 RA Passos-Bueno M.R.:  
 RT "Collagen XVIII, containing an endogenous inhibitor of angiogenesis  
 RT and tumor growth, plays a critical role in the maintenance of retinal  
 RT structure and in neural tube closure."  
 RL Hum. Mol. Genet. 9:2051-2058(2000).  
 RN [6]  
 RN VARIANT ASN-1437.  
 RX MEDLINE=21518361; PubMed=11606364;  
 RA Tugheiti P., Suzuki O., Godoi P.H., Alves V.A., Sertie A.L.,  
 RA Zorick T., Soares F., Camargo A., Moreira E.S., di Loreto C.,  
 RA Moreira-Filho C.A., Simpson A., Oliva G., Passos-Bueno M.R.:  
 RA "A polymorphism in endostatin, an angiogenesis inhibitor, predisposes  
 RT for the development of prostatic adenocarcinoma."  
 RL Cancer Res. 61:7375-7378(2001).  
 CC -!- FUNCTION: COL18A1 probably plays a major role in determining the  
 CC retinal structure as well as in the closure of the neural tube.  
 CC -!- FUNCTION: Endostatin potentially inhibits endothelial cell  
 CC proliferation and angiogenesis. May inhibit angiogenesis by  
 CC binding to the heparan sulphate proteoglycans involved in growth  
 CC factor signalling.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=alternative splicing; Named isoforms=2;  
 CC Name=Long; Synonyms=NC-493;  
 CC IsoId=P39060-1; Sequence=Displayed;  
 CC Name=Short; Synonyms=NC1-303;  
 CC IsoId=P39060-2; Sequence=VSP\_001155; VSP\_001156;  
 CC in liver, lung and kidney.  
 CC -!- TISSUE SPECIFICITY: Present in multiple organs with highest levels  
 CC in liver, lung and kidney.  
 CC -!- PTM: Prolines at the third position of the tripeptide repeating  
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
 CC -!- POLYMORPHISM: There is an association between a polymorphism in  
 CC position 1437 and prostate cancer. Heterozygous Asn-1437  
 CC individuals have a 2.5 times increased chance of developing  
 CC prostate cancer as compared with homozygous Asp-1437 individuals.  
 CC -!- DISEASE: Defects in COL18A1 are a cause of Knobloch syndrome (KNO)  
 CC [MIM:267750]; an autosomal recessive disorder defined by the  
 CC occurrence of high myopia, vitreoretinal degeneration with retinal  
 CC detachment, macular abnormalities and occipital encephalocele.  
 CC -!- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH  
 CC INTERRUPTED HELICES (FACIT) FAMILY.  
 CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; AF018081; AAC39658.1; --  
 CC EMBL; AF018082; AAC39659.1; --  
 CC EMBL; AL163302; CAB90482.1; --  
 CC EMBL; L22548; AAAS1864.1; --  
 CC EMBL; AF184060; AAF01310.1; ALT\_INT.  
 CC PDB; 1BNL; O2-DEC-98.  
 CC GlycoSuiteDB; P39060; --  
 CC Genew; HGNC:2195; COL18A1.  
 CC MIM; 120328; --  
 CC MIM; 267750; --  
 CC GO; GO:0003581; C:collagen; TAS.  
 CC GO; GO:0007397; P:histogenesis and organogenesis; TAS.  
 CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 CC GO; GO:0007601; P:vision; TAS.  
 CC InterPro; IPR008161; Clg\_helix.  
 CC InterPro; IPR008160; Collagen.  
 CC InterPro; IPR008985; ConA like lec\_gl.  
 CC InterPro; IPR001791; Laminin G.  
 CC Pfam; PF01391; Collagen; 7.  
 CC Pfam; PF02210; TSPN; 1.

DR ProDom; PD000007; Clg\_helix; 1.  
 DR SMART; SM00282; LamG; 1.  
 DR SMART; SM00210; TSPN; 1.  
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;  
 KW Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;  
 KW Polymorphism; 3D-structure.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 1516 COLLAGEN ALPHA 1(XVIII) CHAIN.  
 FT CHAIN 1334 1516 ENDOSTATIN.  
 FT DOMAIN 221 409 TSP N-TERMINAL.  
 FT DOMAIN 410 516 NONHELIICAL REGION 1 (NC1).  
 FT DOMAIN 517 550 TRIPLE-HELIICAL REGION 2 (NC2).  
 FT DOMAIN 551 560 NONHELIICAL REGION 2 (NC2).  
 FT DOMAIN 561 640 TRIPLE-HELIICAL REGION 2 (COL2).  
 FT DOMAIN 641 664 TRIPLE-HELIICAL REGION 3 (NC3).  
 FT DOMAIN 665 786 TRIPLE-HELIICAL REGION 3 (COL3).  
 FT DOMAIN 787 809 NONHELIICAL REGION 4 (NC4).  
 FT DOMAIN 810 892 TRIPLE-HELIICAL REGION 4 (COL4).  
 FT DOMAIN 893 906 NONHELIICAL REGION 5 (NC5).  
 FT DOMAIN 907 948 TRIPLE-HELIICAL REGION 5 (COL5).  
 FT DOMAIN 949 961 NONHELIICAL REGION 6 (NC6).  
 FT DOMAIN 962 1034 TRIPLE-HELIICAL REGION 6 (COL6).  
 FT DOMAIN 1035 1044 NONHELIICAL REGION 7 (NC7).  
 FT DOMAIN 1045 1077 TRIPLE-HELIICAL REGION 7 (COL7).  
 FT DOMAIN 1078 1089 NONHELIICAL REGION 8 (NC8).  
 FT DOMAIN 1090 1111 TRIPLE-HELIICAL REGION 8 (COL8).  
 FT DOMAIN 1112 1118 TRIPLE-HELIICAL REGION 9 (NC9).  
 FT DOMAIN 1119 1173 TRIPLE-HELIICAL REGION 9 (COL9).  
 FT DOMAIN 1174 1186 NONHELIICAL REGION 10 (NC10).  
 FT DOMAIN 1187 1204 TRIPLE-HELIICAL REGION 10 (COL10).  
 FT DOMAIN 1205 1516 NONHELIICAL REGION 11 (NC11).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 144 164 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 691 691 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 1329 1329 O-LINKED (GALNAC. .).  
 FT /FTID-CAR 000150.  
 FT /FTID-VSP 001155.  
 FT BY SIMILARITY.  
 FT DISULFID 1366 1506 CELL ATTACHMENT SITE (POTENTIAL).  
 FT DISULFID 1468 1498 CELL ATTACHMENT SITE (POTENTIAL).  
 FT SITE 1095 1097 MISSING (in isoform Short).  
 FT VARSPLIC 1 180 /FTID-VSP 001155.  
 FT VARSPLIC 181 215 HTTEAGTLPAPTPSPSPSLGRPWAPLTPSPSPSS -> MA  
 FT PRCPWPMPRRRLDLVPLAPLVLLGVRAASAP (in  
 FT isoform Short).  
 FT /FTID-VSP 001155.  
 FT VARIANT 1437 1437 D -> N (increased risk of developing  
 FT prostate cancer).  
 FT /FTID-VAR 012709.  
 FT CONFLICT 428 428 F -> S (IN REF. 2).  
 FT CONFLICT 841 841 I -> V (IN REF. 2).  
 FT CONFLICT 877 877 V -> L (IN REF. 3).  
 FT CONFLICT 886 886 P -> R (IN REF. 3).  
 FT CONFLICT 912 912 P -> R (IN REF. 3).  
 FT CONFLICT 933 933 R -> L (IN REF. 3).  
 FT CONFLICT 975 975 P -> L (IN REF. 3).  
 FT CONFLICT 1064 1064 A -> P (IN REF. 3).  
 FT CONFLICT 1084 1084 L -> K (IN REF. 3).  
 FT CONFLICT 1120 1120 P -> A (IN REF. 3).  
 FT CONFLICT 1123 1123 P -> A (IN REF. 3).  
 FT CONFLICT 1126 1126 P -> PPGP (IN REF. 2).  
 FT CONFLICT 1206 1206 G -> GQ (IN REF. 3).  
 FT CONFLICT 1304 1304 R -> G (IN REF. 3).  
 FT CONFLICT 1314 1314 A -> G (IN REF. 3).  
 FT CONFLICT 1323 1324 LR -> CG (IN REF. 3).  
 FT CONFLICT 1443 1443 R -> T (IN REF. 4).  
 FT CONFLICT 1483 1483 S -> Y (IN REF. 4).  
 SQ SEQUENCE 1516 AA; 153840 MW; 3C70F29A4476EE76 CRC64;

Query Match

Best Local Similarity 100.0%; Score 893; DB 1; Length 1516;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 VALNSPLSGMIRGADQCFOQARAVAGLAGTFRFLSRLQDLYSIVRRADRAAVPIV 60  
Db 1346 VALNSPLSGMIRGADQCFOQARAVAGLAGTFRFLSRLQDLYSIVRRADRAAVPIV 1405  
QY 61 NLKDELFFSWALFSGSGPLKPGRIEFSFGKQVLRHPTWPKQSVHSGDPNGRRLTE 120  
Db 1406 NLKDELFFSWALFSGSGPLKPGRIEFSFGKQVLRHPTWPKQSVHSGDPNGRRLTE 1465  
QY 121 SYCETWRTAPSGATGQASSLLGRLGQASCHHAYIVLCIENSPWTAS 170  
Db 1466 SYCETWRTAPSGATGQASSLLGRLGQASCHHAYIVLCIENSPWTAS 1515

RESULT 2  
CAIH MOUSE  
ID CALH MOUSE STANDARD; PRT: 1774 AA.  
AC P3061: Q60672; Q61437; Q62001; Q62002; Q9JK63;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DE 10-OCT-2003 (Rel. 42, Last annotation update)  
GN Collagen alpha 1(XVIII) chain precursor [Contains: Endostatin].  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RC STRAIN=BALB/C; TISSUE=Liver;  
RX MEDLINE=94245707; PubMed=818673;  
RA Rehn M.V., Hintikka E., Pihlajaniemi T.;  
RT "Primary structure of the alpha 1 chain of mouse type XVIII collagen,  
RT partial structure of the corresponding gene, and comparison of the  
RT alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen  
RT chain.";  
RL J. Biol. Chem. 269:13929-13935 (1994).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RX MEDLINE=96435922; PubMed=8838808;  
RA Rehn M., Hintikka E., Pihlajaniemi T.;  
RT "Characterization of the mouse gene for the alpha-1 chain of type  
RT XVIII collagen (COL18A1) reveals that the three variant N-terminal  
RT polypeptide forms are transcribed from two widely separated  
RT promoters.";  
RL Genomics 32:436-446 (1996).  
RN [3]  
RP SEQUENCE OF 1-1387 FROM N.A. (ISOFORM 3).  
RX MEDLINE=94240112; PubMed=8103894;  
RA Rehn M.V., Pihlajaniemi T.;  
RT "Alpha 1(XVIII), a collagen chain with frequent interruptions in the  
RT collagenous sequence, a distinct tissue distribution, and homology  
RT with type XV collagen.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238 (1994).  
RN [4]  
RP SEQUENCE OF 487-1774 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=94240111; PubMed=8183893;  
RA Oh S.P., Kamagata Y., Muragaki Y., Timmons S., Ooshima A., Olsen B.R.;  
RT "Isolation and sequencing of cDNAs for proteins with multiple domains  
RT of Gly-Xaa-Yaa repeats identify a distinct family of collagenous  
RT proteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:4229-4233 (1994).  
RN [5]  
RP SEQUENCE OF 1591-1774 FROM N.A.  
RX MEDLINE=21217748; PubMed=11321448;  
RA Jia S., Zhu F., Li H., He F., Xiu R.-J.;  
RT "Anticancer treatment of endostatin gene therapy by targeting tumor  
RT neovasculature in C57/BL mice.";  
RL Clin. Hemorheol. Microcirc. 23:251-257 (2000).  
RN [6]  
RP CHARACTERIZATION OF ENDOSTATIN AND PARTIAL SEQUENCE.  
RX MEDLINE=97160848; PubMed=9008168;  
RA O'Reilly M.S., Boehm T., Shing Y., Fukai N., Vasios G., Lane W.S.,

FLynn E., Birkhead J.R., Olsen B.R., Folkman J.;  
RT "Endostatin: an endogenous inhibitor of angiogenesis and tumor  
RT growth.";  
RL Cell 88:277-285 (1997).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF ENDOSTATIN.  
RX MEDLINE=98169382; PubMed=9501087;  
RA Hohenester E., Sasaki T., Olsen B.R., Timpl R.;  
RT "Crystal structure of the angiogenesis inhibitor endostatin at 1.5-A  
RT resolution.";  
RL EMBO J. 17:1656-1664 (1998).  
CC -!- FUNCTION: Endostatin potently inhibits endothelial cell  
CC proliferation and angiogenesis. May inhibit angiogenesis by  
CC binding to the heparan sulfate proteoglycans involved in growth  
CC factor signaling.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative promoter;  
CC Comment=2 isoforms, 1 (shown here) and 3, are produced by use  
CC of alternative promoters;  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1; Synonyms=NCI-764;  
CC IsoId=P39061-3; Sequence=Displayed;  
CC Name=2; Synonyms=Long, NCI-517;  
CC IsoId=P39061-1; Sequence=VSP 008303;  
CC Note=Produced by alternative splicing of isoform 1;  
CC Name=3; Synonyms=Short, NCI-301;  
CC IsoId=P39061-2; Sequence=VSP 001157, VSP 001158;  
CC -!- PTM: Prolines at the third position of the tripeptide repeating  
CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
CC -!- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH  
CC INTERRUPTED HELICES (FACIT) FAMILY.  
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.  
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L16898; AAA37434.1; -  
DR EMBL; U03714; AAA20657.1; -  
DR EMBL; U03715; AAC52901.1; -  
DR EMBL; U34606; AAC52901.1; JOINED.  
DR EMBL; U34608; AAC52901.1; JOINED.  
DR EMBL; U34609; AAC52901.1; JOINED.  
DR EMBL; U34610; AAC52901.1; JOINED.  
DR EMBL; U34611; AAC52901.1; JOINED.  
DR EMBL; U34612; AAC52901.1; JOINED.  
DR EMBL; U34613; AAC52901.1; JOINED.  
DR EMBL; U03716; AAC52901.1; JOINED.  
DR EMBL; U03718; AAC52901.1; JOINED.  
DR EMBL; U03715; AAC52902.1; -  
DR EMBL; U34607; AAC52902.1; JOINED.  
DR EMBL; U34608; AAC52902.1; JOINED.  
DR EMBL; U34609; AAC52902.1; JOINED.  
DR EMBL; U34610; AAC52902.1; JOINED.  
DR EMBL; U34611; AAC52902.1; JOINED.  
DR EMBL; U34612; AAC52902.1; JOINED.  
DR EMBL; U34613; AAC52902.1; JOINED.  
DR EMBL; U03716; AAC52902.1; JOINED.  
DR EMBL; U03718; AAC52902.1; JOINED.  
DR EMBL; U03715; AAC52903.1; -  
DR EMBL; U03716; AAC52903.1; JOINED.  
DR EMBL; U03718; AAC52903.1; JOINED.  
DR EMBL; U34607; AAC52903.1; JOINED.  
DR EMBL; U34608; AAC52903.1; JOINED.  
DR EMBL; U34609; AAC52903.1; JOINED.  
DR EMBL; U34610; AAC52903.1; JOINED.  
DR EMBL; U34611; AAC52903.1; JOINED.  
DR EMBL; U34612; AAC52903.1; JOINED.

DR EMBL; U34613; AAC52903.1; JOINED.  
 DR EMBL; U11636; AAC52178.1; --  
 DR EMBL; U11637; AAC52179.1; --  
 DR EMBL; U22545; AAA19787.1; --  
 DR EMBL; AF257775; AAF69009.1; --  
 DR FIR; A56101; A56101.  
 DR PDB; 1KOE; 16-FEB-99.  
 DR PDB; 1DXO; 11-APR-00.  
 DR PDB; 1DX1; 21-JAN-01.  
 DR MGD; MGI:88451; Coll8a1.  
 DR GO; GO:0005604; C:basement membrane; IDA.  
 DR GO; GO:0001525; P:angiogenesis; IMP.  
 DR InterPro; IPR008161; Clg helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008985; ConA like lec\_gl.  
 DR InterPro; IPR001791; Laminin\_G.  
 DR InterPro; IPR003129; TSPN.  
 DR Pfam; PF01391; Collagen; 8.  
 DR Pfam; PF02210; TSPN; 1.  
 DR ProDom; PD000007; Clg helix; 1.  
 DR SMART; SM00282; LamG; 1.  
 DR SMART; SM00210; TSPN; 1.  
 DR PROSITE; PS00308; FZ; 1.  
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;  
 KW Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;  
 KW Alternative promoter usage; 3D-structure.  
 FT SIGNAL 1 26  
 FT CHAIN 27 1774  
 FT COLLAGEN ALPHA 1 (XVIII) CHAIN.  
 FT CHAIN 1591 1774  
 FT ENOSTATIN.  
 FT DOMAIN 245 433  
 FT TSP N-TERMINAL.  
 FT DOMAIN 365 482  
 FT FZ.  
 FT DOMAIN 27 785  
 FT NONHELIICAL REGION 1 (NC1).  
 FT TRIPLE-HELICAL REGION 1 (COL1).  
 FT DOMAIN 813 822  
 FT NONHELIICAL REGION 2 (NC2).  
 FT TRIPLE-HELICAL REGION 2 (COL2).  
 FT DOMAIN 823 896  
 FT NONHELIICAL REGION 3 (NC3).  
 FT TRIPLE-HELICAL REGION 3 (COL3).  
 FT DOMAIN 897 920  
 FT NONHELIICAL REGION 4 (NC4).  
 FT TRIPLE-HELICAL REGION 4 (COL4).  
 FT DOMAIN 921 1042  
 FT NONHELIICAL REGION 5 (NC5).  
 FT TRIPLE-HELICAL REGION 5 (COL5).  
 FT DOMAIN 1043 1065  
 FT NONHELIICAL REGION 6 (NC6).  
 FT TRIPLE-HELICAL REGION 6 (COL6).  
 FT DOMAIN 1066 1148  
 FT NONHELIICAL REGION 7 (NC7).  
 FT TRIPLE-HELICAL REGION 7 (COL7).  
 FT DOMAIN 1149 1162  
 FT NONHELIICAL REGION 8 (NC8).  
 FT TRIPLE-HELICAL REGION 8 (COL8).  
 FT DOMAIN 1163 1204  
 FT NONHELIICAL REGION 9 (NC9).  
 FT TRIPLE-HELICAL REGION 9 (COL9).  
 FT DOMAIN 1205 1217  
 FT NONHELIICAL REGION 10 (NC10).  
 FT TRIPLE-HELICAL REGION 10 (COL10).  
 FT DOMAIN 1218 1290  
 FT NONHELIICAL REGION 11 (NC11).  
 FT TRIPLE-HELICAL REGION 11 (COL11).  
 FT DOMAIN 1291 1300  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DOMAIN 1301 1333  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DOMAIN 1334 1345  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DOMAIN 1346 1369  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DOMAIN 1370 1376  
 FT CELL ATTACHMENT SITE (POTENTIAL).  
 FT DOMAIN 1377 1428  
 FT Missing (in isoform 3).  
 FT DOMAIN 1429 1441  
 FT /Ftid=VSP\_001157.  
 FT DOMAIN 1442 1459  
 FT AGRLPVVCSLPSQEDGCVFTGPA -> MAPRWHLLDV  
 FT LSLVLLVAVRSWAEP (in isoform 3).  
 FT DOMAIN 1460 1774  
 FT /Ftid=VSP\_001158.  
 FT Missing (in isoform 2).  
 FT CARBOHYD 354 354  
 FT /Ftid=VSP\_008303.  
 FT CARBOHYD 361 361  
 FT Missing (in isoform 4).  
 FT CARBOHYD 585 585  
 FT P -> L (IN REF. 4).  
 FT CARBOHYD 947 947  
 FT P -> F (IN REF. 4).  
 FT DISULFID 1623 1763  
 FT A -> R (IN REF. 4).  
 FT DISULFID 1725 1755  
 FT SITE 1351 1353  
 FT VARSPLIC 1 459  
 FT VARSPLIC 460 486  
 FT VARSPLIC 240 486  
 FT VARSPLIC 1147 1147  
 FT CONFLICT 1194 1194  
 FT CONFLICT 1211 1211  
 FT CONFLICT 1211 1211

Query Match 87.1%; Score 778; DB 1; Length 1774;  
 Best Local Similarity 85.8%; Pred. No. 5.4e-69;  
 Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

1 VALNSPLSGMEGIRGADPQCFQARAVGLCTAFSSRLQDLYSIVRRADRAVPIV 60  
 DB VALNTPLSGMGRIGADPQCFQARAVGLSTFFRAFLSSRLQDLYSIVRRADRGVPIV 1662  
 QY 61 NLKDELLFSPWALFSGSGEPLKPGARIPFSGKDVLRHTWPQKSVWHGSDPNGRRLTE 120  
 DB NLKDEVLSFSDLSFSGSQGLPQCARIFSPGDRDLRHPAMPQKSVWHGSDPNGRRLME 1722  
 QY 121 SYCETWRTTEAPSATQASLLGGRLLGQSAASHAYIVLCIENSFMTA 169  
 DB 1723 SYCETWRTTGTATGQASLLSRLLEQKAASHNSIYVLCIENSFMTS 1771

RESULT 3  
 CALE\_HUMAN  
 ID CALE\_HUMAN STANDARD; PRT; 1388 AA.  
 AC P39059;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Collagen alpha 1(XV) chain precursor.  
 GN COL15A1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Umbilical cord;  
 RX MEDLINE=94148920; PubMed=8106446;  
 RA Kivirikko S., Heinamaki P., Rehn M.V., Honkanen N., Myers J.C.,  
 RA Pihlajaniemi T.;  
 RT "Primary structure of the alpha 1 chain of human type XV collagen and  
 RT exon-intron organization in the 3' region of the corresponding  
 RT gene";  
 RL J. Biol. Chem. 269:4773-4779(1994).  
 RN [2]  
 RP SEQUENCE OF 1-569 FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=94140817; PubMed=8307960;  
 RA Murgaki Y., Abe N., Ninomiya Y., Olsen B.R., Ooshima A.;  
 RT "The human alpha 1(XV) collagen chain contains a large amino-terminal  
 RT non-triple helical domain with a tandem repeat structure and homology  
 RT to alpha 1(XVIII) collagen";  
 RL J. Biol. Chem. 269:4042-4046(1994).  
 RN [3]  
 RP SEQUENCE OF 544-1252 FROM N.A.  
 RX MEDLINE=93066196; PubMed=1279671;  
 RA Myers J.C., Kivirikko S., Gordon M.K., Dion A.S., Pihlajaniemi T.;  
 RT "Identification of a previously unknown human collagen chain, alpha  
 RT 1(XV), characterized by extensive interruptions in the triple-helical  
 RT region";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10144-10148(1992).  
 RL TISSUE SPECIFICITY: Expressed predominantly in internal organs  
 CC such as adrenal gland, pancreas and kidney.  
 CC -!- PTM: Prolines at the third position of the tripeptide repeating  
 CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
 CC -!- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH  
 CC INTERRUPTED HELICES (FACIT) FAMILY.  
 CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC EMBL; L25286; AAA58429.1; --  
 DR EMBL; D21230; BAA04762.1; --  
 DR EMBL; L01697; -; NOT\_ANNOTATED\_CDS.

```

DR PIR: A53317; A53317.
DR HSSP: P39061; 1KOE.
DR Genew; KNC:2192; COL15A1.
DR MIM; 120325; -.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 5.
DR Pfam; PF02210; TSPN; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Cell adhesion; Collagen; Glycoprotein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 1388
FT DOMAIN 40 228
FT DOMAIN 229 555
FT DOMAIN 556 573
FT DOMAIN 574 618
FT DOMAIN 619 732
FT DOMAIN 733 763
FT DOMAIN 764 798
FT DOMAIN 799 822
FT DOMAIN 823 867
FT DOMAIN 868 878
FT DOMAIN 879 949
FT DOMAIN 950 983
FT DOMAIN 984 1013
FT DOMAIN 1014 1027
FT DOMAIN 1028 1045
FT DOMAIN 1046 1052
FT DOMAIN 1053 1107
FT DOMAIN 1108 1117
FT DOMAIN 1118 1132
FT DOMAIN 1133 1388
FT DOMAIN 1388 1388
FT REPEAT 358 408
FT REPEAT 409 459
FT REPEAT 460 509
FT REPEAT 510 555
FT CARBOHYD 306 306
FT CARBOHYD 324 324
FT CARBOHYD 687 687
FT CARBOHYD 807 807
FT CARBOHYD 814 814
FT CARBOHYD 1046 1046
FT CONFLICT 10 10
FT CONFLICT 49 49
FT CONFLICT 95 95
FT CONFLICT 150 150
FT CONFLICT 204 204
FT CONFLICT 409 409
SQ SEQUENCE 1388 AA; 141930 MW; 60822AD925A3093D CRC64;

Query Match
Best Local Similarity 56.4%; Score 504; DB 1; Length 1388;
Matches 95; Conservative 27; Mismatches 41; Indels 4; Gaps 1;

Qy 2 ALNPLSGMVGIRGADPQCQARAVGLAGTTFRAFLSSRLQDLYSIVRRADRAAPVIVN 61
Db 1222 ALNPFSGDIR---ADFQCFKQARAGLLSTYRAFLSHLQDLSTIVRKAERYSLPIV 1277

Qy 62 LKDELLFSPSWEALFSGSGPLKPGARIESFGKQVLRHPTWPKSVHSGSPNGRRLTES 121
Db 1278 LKGQVLFNNWDSIFSGHGQGNMHIPIYSFQGRDITMDPSWFQKVIHGGSPHGVRLVDN 1337

Qy 122 YCETWRTEAPATGQASSILGRLGQSAACHAYIVLCIENSPMT 168
Db 1338 YCEAWRTADTAVTGLASPLSTGKILDKQAYCANFLIVLCIENSPMT 1384

RESULT 4

```

---

```

GLI_CHICK
ID GLI_CHICK STANDARD; PRT; 556 AA.
AC P55878;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein GLI1 (GLI) (Fragment).
GN GLI1 OR GLI.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105842; PubMed=8948590;
RA Marigo V., Johnson R.L., Vortkamp A., Tabin C.J.;
RT "Sonic hedgehog differentially regulates expression of GLI and GLI3
RT during limb development.";
RL Dev. Biol. 180:273-283(1996).
CC -I- FUNCTION: MAY REGULATE THE TRANSCRIPTION OF SPECIFIC GENES DURING
CC NORMAL DEVELOPMENT. MAY PLAY A ROLE IN CRANIOFACIAL DEVELOPMENT
CC AND DIGITAL DEVELOPMENT, AS WELL AS DEVELOPMENT OF THE CENTRAL
CC NERVOUS SYSTEM AND GASTROINTESTINAL TRACT. IMPLICATED IN THE
CC TRANSDUCTION OF SHH SIGNAL (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U60762; AAB51659.1; -.
CC HSSP: P08151; 2GLI.
CC InterPro; IPR007087; Znf_C2H2.
CC Pfam; PF00096; zf-C2H2; 5.
CC SMART; SM00355; Znf_FINGER_C2H2_1; 4.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_2; 5.
CC PROSITE; PS01017; ZINC_FINGER_C2H2_2; 5.
CC Zinc-finger; Metal-binding; DNA-binding; Transcription regulation;
CC Nuclear protein; Repeat.
CC ZN_FING 247 272 C2H2-TYPE.
CC ZN_FING 280 307 C2H2-TYPE.
CC ZN_FING 313 337 C2H2-TYPE.
CC ZN_FING 343 368 C2H2-TYPE.
CC ZN_FING 374 399 C2H2-TYPE.
CC ZN_FING 556 556
CC NON_TER 556 556
SQ SEQUENCE 556 AA; 60215 MW; 722D2A5A1CA4D98 CRC64;

Query Match
Best Local Similarity 8.5%; Score 76; DB 1; Length 556;
Best Local Similarity 23.2%; Pred. No. 7.3;
Matches 48; Conservative 21; Mismatches 66; Indels 72; Gaps 10;

Qy 6 PLSG---GMGIRGADPQ-C-----
Db 20 PLHGASAGTFLQGLDLPFVCHQPNLASHHGVLVQTEHPGGADGSRSTPRGAKLG 79
Qy 33 TTRAF-----LSSRLQDLYSIVRRADRAAPVIVN-----LKDELLFSPS----- 70
Db 80 KKRALSTSPISDSDVDLQTVIRTPNSLVAFINSRCASAGSYGHLSTISPSLGYQNP 139
Qy 71 -----WEALFSGSEGPLKPGARIESFGKQVLRHPTWPKSVHSGSPNGRR-----LT 119
Db 140 PGQKQGGQQLFSHTP-PLPFCSSHETLSRRLHPTARGTIKHCQQLKLSLSPLT 198
Qy 120 ESYCETWRTE-----APSATGQASSLLG 142
Db 199 AKYPEE-KSEGDYSSPASTGTQDPLLG 224

```

```

RESULT 5
XA5L_MOUSE
ID XA5L_MOUSE STANDARD; PRT; 334 AA.
AC Q9WTJ8; Q80X84; Q9D993;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE XAP-5-like protein.
GN X5L OR DOH6S2654E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=20005935; PubMed=10534398;
RA Sedlack Z., Muenstermann E., Dhone-Pollet S., Otto C., Bock D.,
RA Schuetz G., Poustka A.;
RT "Human and mouse XAP-5 and XAP-5-like (X5L) genes: identification of
RL an ancient functional retroposon differentially expressed in testis.";
RN Genomics 61:125-132(1999).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okicho T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

```

Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- TISSUE SPECIFICITY: Widely expressed. Abundant in testis, where it  
is expressed in seminiferous tubules, not in the interstitium. At  
the cellular level, expressed in primary spermatocytes and round  
spermatis, but not detectable in spermatogonia, elongating  
spermatis, mature spermatozoa, Sertoli cells or Leydig cells.  
CC -1- SIMILARITY: Belongs to the XAP5 family.  
CC -1- CAUTION: Ref.2 sequence differs from that shown due to a  
frameshift in position 303.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; Y18506; CAB46283.1; -;  
EMBL; Y18508; CAB46284.1; -;  
EMBL; AK007247; BAB24914.1; ALT\_FRAME.  
EMBL; BC049659; AAH49659.1; -;  
MGD; MGI:1351640; DOH6S2654E.  
InterPro: IPR007005; XAP5.  
PFam: PF04921; XAP5; 1.  
CONFLICT 142 142 A -> V (IN REF. 3).  
CONFLICT 229 229 L -> V (IN REF. 2).  
SEQUENCE 334 AA; 39594 MW; 4A438C52BC4A97C6 CRC64;

Query Match 8.5%; Score 75.5; DB 1; Length 334;  
Best Local Similarity 24.8%; Pred. No. 4.5;  
Matches 35; Conservative 13; Mismatches 38; Indels 55; Gaps 7;

QY 8 SCGMGIRGADFCQQAARAVGLAGTFAFLSSRLQDLYSIVRRADRAAV-PIVNLKDEL 66  
DB 204 SGHRTVR-----MSKGSTVQOFLKRALQGLRDFELRAAGVEQLMYVKEDL 251  
QY 67 LPFSNEALFS-----GSEGLKPGARFSD-----GKDVLRH 99  
DB 252 ILPHVHTFDFTVAKARGSGFL-----FSFDVDDVRLSDATWKEDESHAGKVLLR- 304

QY 100 PTWPQK-----SVMHGSDP 113  
DB 305 -SWYEKXKHIFPASRWEPYDP 324

RESULT 6  
Y4BG RHISN STANDARD; PRT; 271 AA.  
AC P55374;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical protein Y4BG precursor.  
GN Y4BG.  
OS Rhizobium sp. (strain NGR234).  
OG Plasmid sym pNGR234a.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
OX NCBI\_TaxID=394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97305956; PubMed=9163424;  
RA Freiberg C.A., Felly R., Bairoch A., Broughton W.J., Rosenthal A.,  
RA Perret X.;  
RT "Molecular basis of symbiosis between Rhizobium and legumes.";

```

RL Nature 387:394-401(1997).
CC -!- SIMILARITY: None obvious.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB000066; AAB91622.1; --
CC DR HYPOTHETICAL protein; Plasmid; Signal.
CC KW SIGNAL 1 22 POTENTIAL.
CC FT CHAIN 23 271 HYPOTHETICAL PROTEIN Y4BG
CC SQ SEQUENCE 271 AA; 30034 MW; 5572CF1461FCD463 CRC64;

Query Match      8.4%; Score 75; DB 1; Length 271;
Best Local Similarity 21.1%; Pred. No. 4;
Matches 28; Conservative 26; Mismatches 35; Indels 44; Gaps 7;

QY 4 NSPLSGMGRGIGADFCQFOQARAVGLACTFAFLSSRLQDLYSIVRRA-----DRAAV 57
DB 174 NADIAATIKSLGADGFEAFQALITGSGEFPKS-----DDY-IGRTCTPHMCREQAL 225
QY 58 PTVNLKDELLPSPSEALFSGSEGLPKPGARIFSDGKDVLRHP---TWPKSVWHGSDPN 114
DB 226 LFLSAKDRAYAAW-----XP-----HKKLIIVHPVKQWPEKA----- 259
QY 115 GRRLTESYCETWR 127
DB 260 -KQELRAWAETWK 271

RESULT 7
CRK_XENLA
ID CRK_XENLA STANDARD; PRT; 296 AA.
AC P87378;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE SH2/SH3 adaptor crk (Adapter molecule crk) (CRK2).
GN CRK.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97180926; PubMed=9029144;
RA Evans E.K., Lu W., Strum S.L., Mayer B.J., Kornbluth S.;
RT "Crk is required for apoptosis in Xenopus egg extracts.";
RL EMBO J. 16:230-241(1997).
CC -!- FUNCTION: Required for apoptosis in Xenopus egg extracts. May
CC mediate attachment-induced MAPK8 activation, membrane ruffling and
CC cell motility in a Rac-dependent manner.
CC -!- SUBUNIT: Interacts with ABL, C3G, SOS, MAP4K1, MAPK8 and DOCK1 via
CC its first SH3 domain. Interacts with BCAR1, CBL, PXN and GAB1 via
CC its SH2 domain upon stimulus-induced tyrosine phosphorylation.
CC Interacts with several tyrosine-phosphorylated growth factor
CC receptors such as EGFR, PDGFR and INSR via its SH2 domain (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated to the plasma
CC membrane upon cell adhesion (By similarity).
CC -!- PTM: Phosphorylated on Tyr-214 upon cell adhesion. Results in the
CC negative regulation of the association with SH2- and SH3-binding
CC partners, possibly by the formation of an intramolecular
CC interaction of phosphorylated Tyr-221 with the SH2 domain. This
CC leads finally to the down-regulation of the Crk signaling pathway
CC (By similarity).
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -!- SIMILARITY: Contains 2 SH3 domains.

```

---

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U89774; AAB49698.1; --
CC DR HSPF; Q64010; ICKA.
CC DR InterPro; IPR000980; SH2.
CC DR InterPro; IPR001452; SH3.
CC DR Pfam; PF00017; SH2; 1.
CC DR Pfam; PF00018; SH3; 1.
CC DR PRINTS; PR00401; SH2DOMAIN.
CC DR PRINTS; PR00452; SH3DOMAIN.
CC DR ProDom; PD000093; SH2; 1.
CC DR ProDom; PD000066; SH3; 1.
CC DR SMART; SM00252; SH2; 1.
CC DR SMART; SM00326; SH3; 2.
CC DR PROSITE; PS50001; SH2; 1.
CC DR PROSITE; PS50002; SH3; 1.
CC KW SH2 domain; SH3 domain; Repeat; Phosphorylation.
FT DOMAIN 13 112
FT DOMAIN 125 185
FT DOMAIN 249 289
FT MOD_RES 214 214 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 296 AA; 33409 MW; 544F11F41F75A66 CRC64;

Query Match      8.3%; Score 74; DB 1; Length 296;
Best Local Similarity 32.4%; Pred. No. 5;
Matches 24; Conservative 12; Mismatches 20; Indels 18; Gaps 5;

QY 75 FSGSEG---PLKPGARIFSDGKDVLRHP---TWPKSVWHGSDPNRR--LTESYCETWTE 129
DB 136 FIGNDEDLFPKKG-----DILIRKPEQWNAEDSDGRGMIPVPYVKYR-- 184
QY 130 APSATGQASSLLGG 143
DB 185 PPSPPG--SALLGG 196

RESULT 8
STRN_RAT
ID STRN_RAT STANDARD; PRT; 780 AA.
AC P70483;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Striatin.
GN STRN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 222-243; 258-269; 451-456; 488-494
RP AND 764-777.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=96354919; PubMed=8769426;
RA Castets F., Bartoli M., Barnier J.V., Baillat G., Salin P.,
RA Mogrich A., Bourgeois J.-P., Denizot F., Rougon G., Calothy G.,
RA Monneron A.;
RT "A novel calmodulin-binding protein, belonging to the WD-repeat
RT family, is localized in dendrites of a subset of CNS neurons.";
RL J. Cell Biol. 134:1051-1062(1996).
CC -!- FUNCTION: Binds calmodulin in a calcium dependent manner. May
CC function as scaffolding or signaling protein.
CC -!- SUBUNIT: INTERACTS WITH PROTEIN PHOSPHATASE 2A (PP2A) (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-bound.
CC -!- TISSUE SPECIFICITY: MAINLY EXPRESSED IN THE CENTRAL NERVOUS
CC SYSTEM. MOSTLY CONFINED IN DENDRITES, NOT IN AXONS, AND IS MOST

```







CC WITH POLYPHOSPHATE AS SUBSTRATE.  
 CC -!- CATALYTIC ACTIVITY: {polyphosphate} (N) + H(2)O =  
 CC {polyphosphate} (N-1) + phosphate.  
 CC -!- COFACTOR: Magnesium.  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SUBCELLULAR LOCATION: Membrane-associated.  
 CC -!- SIMILARITY: Belongs to the gppA / ppx family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; L06129; AAA24415.1; -;  
 DR EMBL; AE000336; AAC75595.1; -;  
 DR EMBL; D90878; BAA16390.1; -;  
 DR EMBL; D90880; BAA16392.1; -;  
 DR EMBL; AE016764; AAN81470.1; -;  
 DR EMBL; AE005479; AAG57612.1; -;  
 DR EMBL; AP002561; BAB36787.1; -;  
 DR PIR; A45333; A45333.  
 DR PIR; D91049; D91049.  
 DR PIR; H85893; H85893.  
 DR EcoGene; EG11403; ppx.  
 DR InterPro; IPR003695; Ppx GppA.  
 DR Pfam; PF02541; Ppx-GppA; 1.  
 KW Hydrolase; Magnesium; Membrane; Complete proteome.  
 FT INIT\_MET 0  
 SQ SEQUENCE 512 AA; 58004 MW; 48611AFF5D9FB9C3 CRC64;  
 Query Match 8.2%; Score 73; DB 1; Length 512;  
 Best Local Similarity 21.9%; Pred. No. 13;  
 Matches 46; Conservative 28; Mismatches 66; Indels 70; Gaps 11;  
 QY 22 FQAR-----AVGLAGTFRA-----FLSSRLQDLY 46  
 Db 186 FQARMAAAQKLETLTWQFRIQGNVAMGASGITKAAHEVLMEGKDGITPERLEKLV 245  
 QY 47 -STVRADRAAPVIVNLKDE--LLFSPWEALFSG-----SEGPKPGARIPSF 91  
 Db 246 KEVLIRNFNLSLPLGSEERKTVFVPLGLAILCGVFDALAIRLSDGALRGV-LYEM 304  
 QY 92 DGK----DVLRHPTWPQKSVVHSGSDPNRRL---TESYCTWTEAPS-ATGQASSILG- 142  
 Db 305 EGRFRHQDVSRTASSIANQYHIDSEQARFVLTWTQMYEQWREQPKLAHPQLEALLRW 364  
 QY 143 -----GRLLGQSAASCHAYIVLCIENS 165  
 Db 365 AAMLHEVGLNINSHGLRHSAYI---LQNS 391  
 RESULT 12  
 YP67\_MYCTU STANDARD; PRT; 884 AA.  
 AC Q50654; Q50731;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical protein rv2567/M2843.  
 GN RV2567 OR M2643 OR MTCY227.34C OR MTCY9C4.01C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Broach R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RX MEDLINE=22206494; PubMed=12218036;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
 Delcher A., Ufferback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";  
 RL J. Bacteriol. 184:5479-5490 (2002).  
 CC -!- SIMILARITY: SOME, TO M.TUBERCULOSIS RV2411C AND SYNECHOCYSTIS PCC  
 CC 6803 SLL0335.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; Z77250; CAB01053.1; -;  
 DR EMBL; A5007098; AAK46956.1; -;  
 DR PIR; C70729; C70729.  
 DR TIGR; MT2643; -;  
 DR Tuberculist; RV2567; -;  
 DR InterPro; IPR007296; DUF403.  
 DR InterPro; IPR007297; DUF404.  
 DR InterPro; IPR007302; DUF407.  
 DR Pfam; PF04168; DUF403; 1.  
 DR Pfam; PF04169; DUF404; 1.  
 DR Pfam; PF04174; DUF407; 1.  
 KW Hypothetical protein; Complete proteome.  
 FT CONFLICT 645 645 Q -> R (IN REF. 2).  
 SQ SEQUENCE 884 AA; 95448 MW; 95D23A4D2EDB365 CRC64;  
 Query Match 8.1%; Score 72.5; DB 1; Length 884;  
 Best Local Similarity 29.5%; Pred. No. 28;  
 Matches 31; Conservative 15; Mismatches 26; Indels 33; Gaps 6;  
 QY 51 RADRAAPVIVNLKDELLFSPWEALFSGSGPLKPGARIPFSG-----KDVLRHPTW 102  
 Db 609 RADMTAVA-----PS--TLWSLTVDPRPSGLSVQSGVGLAAQAQVRLQSLNDTW 656  
 QY 103 -----PKSVHSGSDPNRRLTESYCTWTEAPSATGQASSILG 142  
 Db 657 MYLANVERAVEHKSDP-----PQSLAE---ADAVLASAQAEITLAG 693  
 RESULT 13  
 YP97\_MYCBO STANDARD; PRT; 884 AA.  
 AC P59974;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hypothetical protein Mb2597.  
 GN Mb2597.  
 OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1765;

ASP-150.  
MEDLINE=96278913; PubMed=8662307;  
RA Matsuda M., Ota S., Tanimura R., Nakamura K., Takenawa T.,  
RA Nagashima K., Kurata T.;  
RT "Interaction between the amino-terminal SH3 domain of Crk and its  
RT natural target proteins.";  
RT J. Biol. Chem. 271:14468-14472(1996).  
RN [5].  
RP INTERACTION WITH DOCK4.  
RX MEDLINE=22515525; PubMed=12628187;  
RA Yajnik V., Paulding C., Sordella R., McClatchey A.I., Saito M.,  
RA Wahrer D.C.R., Reynolds P., Bell D.W., Lake R., van den Heuvel S.,  
RA Settleman J., Haber D.A.;  
RT "DOCK4, a GTPase activator, is disrupted during tumorigenesis.";  
RL Cell 112:673-684(2003).  
CC -I- FUNCTION: The Crk-I and Crk-II forms differ in their biological  
CC activities. Crk-II has less transforming activity than Crk-I. Crk-  
CC II mediates attachment-induced MAPK8 activation, membrane ruffling  
CC and cell motility in a Rac-dependent manner. Involved in  
CC phagocytosis of apoptotic cells and cell motility via its  
CC interaction with DOCK1 and DOCK4.  
CC -I- SUBUNIT: Interacts with ABL, C3G, SOS, MAP4K1, MAPK8 and DOCK3 via  
CC its first SH3 domain. Interacts with BCAR1, CSL, PXN and GRB1 via  
CC its SH2 domain upon stimulus-induced tyrosine phosphorylation.  
CC Interacts with several tyrosine-phosphorylated growth factor  
CC receptors such as EGFR, PDGFR and INSR via its SH2 domain (By  
CC similarity). Interacts with DOCK1 and DOCK4.  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic; translocated to the plasma  
CC membrane upon cell adhesion (By similarity).  
CC -I- ALTERNATIVE PRODUCTS:  
CC -I- Event=Alternative splicing; Named isoforms=2;  
CC Name=Crk-II;  
CC IsoId=P46108-1; Sequence=Displayed;  
CC Name=Crk-I;  
CC IsoId=P46108-2; Sequence=VSP\_004173;  
CC -I- DOMAIN: The C-terminal SH3 domain function as a negative modulator  
CC for transformation and the N-terminal SH3 domain appears to  
CC function as a positive regulator for transformation (By  
CC similarity).  
CC -I- PTM: Phosphorylation of Crk-II (40 kDa) gives rise to a 42 kDa  
CC form.  
CC -I- PTM: Phosphorylated on Tyr-221 upon cell adhesion. Results in the  
CC negative regulation of the association with SH2- and SH3-binding  
CC partners, possibly by the formation of an intramolecular  
CC interaction of phosphorylated Tyr-221 with the SH2 domain. This  
CC leads finally to the down-regulation of the Crk signaling pathway  
CC (By similarity).  
CC -I- SIMILARITY: Contains 1 SH2 domain.  
CC -I- SIMILARITY: Contains 2 SH3 domains.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; D10656; BAA01505.1; -;  
CC EMBL; S65701; AAB28213.1; -;  
CC FIC; A45022; A45022.  
CC PDB; 1JUS; 06-NOV-02.  
CC SWISS-2DPAGE; P46108; HUMAN.  
CC Genew; HGNC:2362; CRK.  
CC MIM; 164762; -;  
CC GO; GO:0005737; C:cytoplasm; TAS.  
CC GO; GO:0005634; C:nucleus; TAS.  
CC GO; GO:0030036; P:actin cytoskeleton organization and biogenesis; TAS.  
CC GO; GO:0006357; P:regulation of transcription from Pol II promoter; TAS.  
CC GO; GO:0007165; P:signal transduction; TAS.  
CC InterPro; IPR000980; SH2.  
CC Pfam; PF00017; SH2; 1.  
CC





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 08:14:01 ; Search time 42 Seconds  
(without alignments)  
1277.097 Million cell updates/sec

Title: US-09-171-607A-1  
Perfect score: 893  
Sequence: 1 VALNSPLSGMGRGIRGADFQ.....ASCHHAVIVLCIENSFMTAS 170

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:  
1: sp\_archea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	893	100.0	187	Q8WXI5	Q8WXI5 homo sapien
2	893	100.0	261	Q8NG19	Q8NG19 homo sapien
3	893	100.0	816	Q8N4S4	Q8N4S4 homo sapien
4	778	87.1	1140	Q61434	Q61434 mus musculus
5	763	85.4	226	Q1032D	Q1032D rattus norv
6	757	84.8	171	Q8WUWS	Q8WUWS rattus norv
7	722	80.9	160	Q1032D	Q1032D mus musculus
8	711	79.6	1344	Q1032D	Q1032D mus musculus
9	654	73.2	1315	Q8QH19	Q8QH19 gallus gall
10	643	72.0	1307	Q8JFF7	Q8JFF7 xenopus lae
11	606	67.9	361	Q8AWC6	Q8AWC6 brachydanio
12	529	59.2	102	Q86T70	Q86T70 homo sapien
13	515	57.7	1367	Q1032D	Q1032D mus musculus
14	512	57.3	1367	Q1032D	Q1032D mus musculus
15	505	56.6	1388	Q9Y4W4	Q9Y4W4 homo sapien
16	427.5	47.9	950	Q86SC8	Q86SC8 ciona intes

17	424.5	47.5	299	5	Q8MSE3	Q8MSE3 drosophila
18	418.5	46.9	778	5	Q86BH1	Q86BH1 drosophila
19	418.5	46.9	792	5	Q8MT89	Q8MT89 drosophila
20	357	40.0	650	5	O17866	O17866 caenorhabdi
21	357	40.0	778	5	Q9U9K6	Q9U9K6 caenorhabdi
22	357	40.0	1117	5	Q9U9K7	Q9U9K7 caenorhabdi
23	357	40.0	1154	5	Q810G6	Q810G6 caenorhabdi
24	89	10.0	287	16	Q89138	Q89138 bradyrhizob
25	87.5	9.8	1257	16	Q7WF24	Q7WF24 bordetella
26	87.5	9.8	1257	16	Q7W3P8	Q7W3P8 bordetella
27	87.5	9.8	1257	16	Q7VTE8	Q7VTE8 bordetella
28	85	9.5	478	4	Q81VY1	Q81VY1 homo sapien
29	85	9.5	1649	4	Q81VY7	Q81VY7 homo sapien
30	82.5	9.2	208	16	Q92K28	Q92K28 rhizobium m
31	82	9.2	498	16	Q9KXK2	Q9KXK2 streptomyce
32	80.5	9.0	7716	16	Q7UWZ8	Q7UWZ8 rhodopirell
33	80	9.0	409	16	Q82AG1	Q82AG1 streptomyce
34	80	9.0	493	16	Q98AT9	Q98AT9 rhizobium l
35	80	9.0	651	5	Q9VFA9	Q9VFA9 drosophila
36	80	9.0	1024	16	Q8F8H1	Q8F8H1 corynebacte
37	79.5	8.9	240	16	Q8F863	Q8F863 leptospira
38	79.5	8.9	716	10	Q8LJG8	Q8LJG8 oryza sativ
39	79	8.8	477	10	Q9SMY7	Q9SMY7 arabidopsis
40	79	8.8	525	10	Q94JL8	Q94JL8 arabidopsis
41	78.5	8.8	285	16	Q98EUS	Q98EUS rhizobium l
42	78	8.7	1715	6	Q9GLM4	Q9GLM4 bos taurus
43	77.5	8.7	395	11	Q9QUP4	Q9QUP4 mus musculu
44	77	8.6	263	5	Q7YZJ8	Q7YZJ8 hexamita in
45	77	8.6	314	16	Q8U8N8	Q8U8N8 agrobacteri

## ALIGNMENTS

## RESULT 1

Q8WXI5 PRELIMINARY; PRT; 187 AA.

AC Q8WXI5;  
DT 01-MAR-2002 (TREMREL. 20, Created)  
DT 01-MAR-2002 (TREMREL. 20, Last sequence update)  
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)  
DE Collagen XVIII (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21409408; PubMed=11517600;  
RA Feng Y., Cui L.B., Liu C.X., Ma Q.J.;  
RT "Inhibition effect in vitro of purified endostatin expressed in Pichia pastoris".  
RL Sheng Wu Gong Cheng Xue Bao 17:278-282(2001).  
DR EMBL; AF416592; AAL37720.1; -.  
FT NON\_TER  
SQ SEQUENCE 187 AA; 20448 MW; 72B1047D85838CD3 CRC64;

Query Match	100.0%;	Score 893;	DB 4;	Length 187;
Best Local Similarity	100.0%;	Pred. No. 2,7e-83;		
Matches 170;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	VALNSPLSGMGRGIRGADFQ	Q8WXI5	Q8WXI5
Db	17	VALNSPLSGMGRGIRGADFQ	Q8WXI5	Q8WXI5
QY	61	NKDELFPSPWEALFSGSEGLKPGARIFSDGKDLRHPTWPKSVWHGSDNGRRLLTE	120	
Db	77	NKDELFPSPWEALFSGSEGLKPGARIFSDGKDLRHPTWPKSVWHGSDNGRRLLTE	136	
QY	121	SYCETWTEAPSATGQSSLLGLGRLGQSAASCHHAVIVLCIENSFMTAS	170	
Db	137	SYCETWTEAPSATGQSSLLGLGRLGQSAASCHHAVIVLCIENSFMTAS	186	

Db 706 NLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNRRLTE 1705

QY 121 SYCETWRTPEAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170

Db 766 SYCETWRTPEAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 815

RESULT 4

Q61434 PRELIMINARY; PRT; 1140 AA.

ID Q61434

AC Q61434

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Collagen (Fragment).

GN COL15A1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Abe N., Muragaki Y., Yoshiocka H., Inoue H., Ninomiya Y.;

RT "Identification of a novel collagen chain represented by extensive

RT interruptions in the triple-helical region.";

RL Cell. Mol. Biol. Res. 196:576-582(1993).

DR EMBL; D17546; BAA04483.1; -

DR PIR; B56101; B56101.

DR HSP; F39061; IKOE.

DR MGD; MGI:88449; Collisal.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR GO; GO:0007155; P:cell adhesion; IEA.

DR InterPro; IPR008161; Clg helix.

DR InterPro; IPR008160; Collagen.

DR InterPro; IPR003129; TSPN.

DR Pfam; PF01391; Collagen; 8.

DR Pfam; PF02210; TSPN; 1.

DR Probom; PD000007; Clg\_helix; 1.

KW Collagen.

FT NON\_TER

SQ SEQUENCE 1 115156 MW; 8BC07B6862B3BDFE CRC64;

Query Match 87.1%; Score 778; DB 11; Length 1140;

Best Local Similarity 85.8%; Pred. No. 1.5e-70;

Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGIRGADFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 60

Db 969 VALNTPLSGMGIRGADFCQFQARAVGLSGTFFRAFLSSRLQDLYSIVRRADRGSPV 1028

QY 61 NLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNRRLTE 120

Db 1029 NLKDEVLSPSWDSLFSGSGQQLQPGARIFSPDGKDLRHPTWPKSVWHGSDPSGRRLME 1088

QY 121 SYCETWRTPEAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTA 169

Db 1089 SYCETWRTPEAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTS 1137

RESULT 5

Q9QZD2 PRELIMINARY; PRT; 226 AA.

ID Q9QZD2

AC Q9QZD2

DT 01-WAY-2000 (TrEMBLrel. 13, Created)

DT 01-WAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Collagen XVIII (Fragment).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

Query Match 100.0%; Score 893; DB 4; Length 261;

Best Local Similarity 100.0%; Pred. No. 4.1e-83;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGIRGADFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 60

Db 91 VALNSPLSGMGIRGADFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 150

QY 61 NLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNRRLTE 120

Db 151 NLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNRRLTE 210

QY 121 SYCETWRTPEAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170

Db 211 SYCETWRTPEAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 260

RESULT 3

Q8N4S4 PRELIMINARY; PRT; 816 AA.

ID Q8N4S4

AC Q8N4S4

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Similar to collagen, type XVIII, alpha 1 (fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC033715; AAH33715.1; -

DR InterPro; IPR008161; Clg helix.

DR InterPro; IPR008160; Collagen.

DR Pfam; PF01391; Collagen; 5.

DR Probom; PD000007; Clg\_helix; 1.

KW Collagen.

FT NON\_TER

SQ SEQUENCE 1 82553 MW; 5D539B2946694F86 CRC64;

Query Match 100.0%; Score 893; DB 4; Length 816;

Best Local Similarity 100.0%; Pred. No. 1.7e-82;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGIRGADFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 60

Db 646 VALNSPLSGMGIRGADFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 705

QY 61 NLKDELLFPSWEALFSGSEGLPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNRRLTE 120

RC STRAIN=Sprague-Dawley;  
 RX MEDLINE=20227226; PubMed=10766159;  
 RA Perletti G., Concari P., Giardini R., Marras E., Piccinini F.,  
 RA Folkman J., Chen L.;  
 RT "Antitumor activity of endostatin against carcinogen-induced rat  
 RL primary mammary tumors";  
 DR Cancer Res. 60:1793-1796(2000).  
 DR EMBL; AF189709; AAF00975.1; -.  
 DR HSSP; P39061; 1KOE.  
 FT NON TER 1  
 SQ SEQUENCE 226 AA; 25350 MW; 38B83C0486C0E949 CRC64;

Query Match 85.4%; Score 763; DB 11; Length 226;  
 Best Local Similarity 84.6%; Pred. No. 6.8e-70;  
 Matches 143; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

QY 1 VALNPLSGMGRGIRGADFCQQAQAVGLAGTFFAFLSSRLQDLYSIVRRADRAAIVP 60  
 Db 55 VALNPLSGMGRGIRGADFCQQAQAVGLAGTFFAFLSSRLQDLYSIVRRADRAAIVP 114

QY 61 NLKDELLFSPSWDLTFLSSGQQLHSGARIFSDGDRVLRHPAMPQKSVWHGSDPSGRRLME 120  
 Db 115 NLKDEVLSPSWDLTFLSSGQQLHSGARIFSDGDRVLRHPAMPQKSVWHGSDPSGRRLME 174

QY 121 SYCETWRTAPSGATGQASSLLGRLGQSAASHHAYIVLICIENSFMTA 169  
 Db 175 SYCETWRTATGVTGQASSLLGRLGQSAASHHAYIVLICIENSFMTS 223

RESULT 6  
 Q9WUW5  
 ID Q9WUW5 PRELIMINARY; PRT; 171 AA.  
 AC Q9WUW5;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Collagen type XVIII, alpha (I) chain (fragment).  
 GN COL18A1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Jia J.D., Bauer M., Eberstpaecher U., Donner P., Schuppan D.;  
 RT "Partial 3'-terminal cDNA sequence of rat collagen XVIII/endostatin";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ236873; CAB44263.1; -.  
 DR HSSP; P39061; 1KOE.  
 FT NON TER 1  
 SQ SEQUENCE 171 AA; 18933 MW; 81BE2BE3FC2C8E72 CRC64;

Query Match 84.8%; Score 757; DB 11; Length 171;  
 Best Local Similarity 85.0%; Pred. No. 2e-69;  
 Matches 142; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

QY 1 VALNPLSGMGRGIRGADFCQQAQAVGLAGTFFAFLSSRLQDLYSIVRRADRAAIVP 60  
 Db 5 VALNPLSGMGRGIRGADFCQQAQAVGLAGTFFAFLSSRLQDLYSIVRRADRAAIVP 64

QY 61 NLKDELLFSPSWDLTFLSSGQQLHSGARIFSDGDRVLRHPAMPQKSVWHGSDPSGRRLME 120  
 Db 65 NLKDEVLSPSWDLTFLSSGQQLHSGARIFSDGDRVLRHPAMPQKSVWHGSDPSGRRLME 124

QY 121 SYCETWRTAPSGATGQASSLLGRLGQSAASHHAYIVLICIENSFMTA 167  
 Db 125 SYCETWRTATGVTGQASSLLGRLGQSAASHHAYIVLICIENSFMTA 171

RESULT 7  
 Q9CRT2  
 ID Q9CRT2 PRELIMINARY; PRT; 160 AA.  
 AC Q9CRT2;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Procollagen, type XVIII, alpha 1 (fragment).  
 GN COL18A1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Kochava T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kuchwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK014292; BAB29249.1; -.  
 DR HSSP; P39061; 1KOE.  
 DR MGD; MGI:88451; Coll18a1.  
 DR GO; GO:0005604; C:basement membrane; IDA.  
 DR GO; GO:0001525; P:angiogenesis; IMP.  
 FT NON TER 1  
 SQ SEQUENCE 160 AA; 17725 MW; 60F853D777C375D2 CRC64;

Query Match 80.9%; Score 722; DB 11; Length 160;  
 Best Local Similarity 85.4%; Pred. No. 6.9e-66;  
 Matches 134; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 13 GIRGADFCQQAQAVGLAGTFFAFLSSRLQDLYSIVRRADRAAIVPVLKDELLFSPSW 72  
 Db 1 GIRGADFCQQAQAVGLAGTFFAFLSSRLQDLYSIVRRADRAAIVPVLKDELLFSPSW 60

QY 73 ALFSGSEGLPKGARIPIFSDGDRVLRHPAMPQKSVWHGSDPSGRRLME 132  
 Db 61 SLFSGSQQLQPGARIFSDGDRVLRHPAMPQKSVWHGSDPSGRRLME 120

QY 133 ATGQASSLLGRLGQSAASHHAYIVLICIENSFMTA 169  
 Db 121 ATGQASSLLGRLGQSAASHHAYIVLICIENSFMTS 157

RESULT 8  
 Q93419  
 ID Q93419 PRELIMINARY; PRT; 1344 AA.  
 AC Q93419;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)



DR	GO; GO:0007155; P:cell adhesion; IEA.
DR	InterPro; IPR008161; Clg helix.
DR	InterPro; IPR008160; Collagen.
DR	InterPro; IPR003129; TSPN.
DR	Pfam; PF01391; Collagen; 7.
DR	Pfam; PF02210; TSPN; 1.
DR	ProDom; PD000007; Clg helix; 1.
DR	SMART; SM00210; TSPN; 1.
KW	Collagen.
SQ	SEQUENCE 1315 AA; 134946 MW; OC56C23D5E058365 CRC64;
Query Match	73.2%; Score 654; DB 13; Length 1315;
Best Local Similarity	73.2%; Pred. No. 9e-58;
Matches 123; Conservative	16; Mismatches 29; Indels 0; Gaps 0;
Qy	1 VALNSPLSGMGIRGADFCQCQAARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAPVIV 60
Dd	1144 VALNAPLSGMKSIKRGVDVDFCFEQARKSLGHGTFFRAFLSSRLQDLYSIVRRADRSQVIV 1203
Qy	61 NLKDLELPPSWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPQKSVWHGSDPNGRLLTE 120
Dd	1204 NLRDEVLYNDWSLFSGSEAQMKGARIFSPDGKDVATDTFWPKMWHGSDAKGRLLTE 1263
Qy	121 SYCETWRTEAPSATGQASSLLGGRLLGQSAASHYIIVLCIENSPMT 168
Dd	1264 SYCETWRTEASVTQSASSLTSGKLLEQRPSCKNFIVLCIENSPMT 1311
RESULT 10	
Q8JUFF7	PRELIMINARY; PRF; 1307 AA.
ID	Q8JUFF7
AC	Q8JUFF7;
DT	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Type XVIII collagen short variant.
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC	Xenopodinae; Xenopus.
RX	NCBI_TaxID=8355;
QN	[1]
RP	SEQUENCE FROM N.A.
RK	MEDLINE=22166979; PubMed=12175494;
RA	Elamra H., Peterson J., Pihlajaniemi T., Destree O.;
RT	"Cloning of three variants of type XVIII collagen and their expression
RT	patterns during Xenopus laevis development.";
RL	Mech. Dev. 114:109-113(2002).
DR	EMBL; AV052763; AAL14257.1;
DR	GO; GO:0005198; F:structural molecule activity; IEA.
DR	GO; GO:0007155; P:cell adhesion; IEA.
DR	InterPro; IPR008160; Collagen.
DR	InterPro; IPR003129; TSPN.
DR	Pfam; PF01391; Collagen; 6.
DR	Pfam; PF02210; TSPN; 1.
DR	SMART; SM00210; TSPN; 1.
KW	Collagen.
SQ	SEQUENCE 1307 AA; 134153 MW; D53EDBF3DE34976 CRC64;
Query Match	72.0%; Score 643; DB 13; Length 1307;
Best Local Similarity	71.4%; Pred. No. 1.2e-56;
Matches 120; Conservative	18; Mismatches 30; Indels 0; Gaps 0;
Qy	1 VALNSPLSGMGIRGADFCQCQAARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAPVIV 60
Dd	1136 VALNAPISGMKSIKRGVDVDFCFEQARKAGLHGHTFFRAFLSSRLQDLYSIVRRADRSQVIV 1195
Qy	61 NLKDLELPPSWALFSGSEGPLKPGARIFSPDGKDVLRHPTWPQKSVWHGSDPNGRLLTE 120
Dd	1196 NLRDEVLYNDWSLFSGSEAQMKGARIIUSPDGKVTTDTPWPKMWHGSDAKGRLLTE 1255
Qy	121 SYCETWRTEAPSATGQASSLLGGRLLGQSAASHYIIVLCIENSPMT 168

```

Db      1256 SYCETWRTDES AVTGQASSLTSKLLQRPQSCNKNFIVLCIENSFMT 1303
RESULT 11
Q8AWC6 PRELIMINARY; PRT; 361 AA.
AC Q8AWC6
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Collagen XVIII (Fragment).
GN COL18A1
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]_TaxID=7955;
RP SEQUENCE FROM N.A.
RA Haftik Z., Morvan-Dubois G., Thisse B., Gartone R., Le Guellec D.;
RT "Sequence and embryonic expression of collagen XVIII NC11 domain
RT (endostatin) in the zebrafish."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ494637; CAD38825.1; -.
FT NON_TER
SQ SEQUENCE 361 AA; 40222 MW; 3C5A0F8479D26735 CRC64;
Query Match 67.9%; Score 606; DB 13; Length 361;
Best Local Similarity 66.5%; Pred. No. 1.4e-53;
Matches 113; Conservative 22; Mismatches 35; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGRGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADRAVPV 60
Db      190 IALNSPQVGNMGRGADFCQFQARAVGMKGTFRFLSSKLQDLYSIVRRSDRETLPV 249
QY 61 NLKDELLFPSEWALFSGSEGLKPGARIFSGDKVLRHPTPQKSVWHGSDPNRRRLTE 120
Db      250 NLKQVLFSSWESLFSDESRLKAPYISFGDVLDRDSAWPEKMIWHGSDGRGHRQTD 309
QY 121 SYCETWRTDEAPTSATQASSLLGRLGQSAASHAYIVLCIENSFMTAS 170
Db      310 NYCETWRAGDRAVTLGLASSLQGLLQQTSSCSSSYIALCIENSYMTQS 359
RESULT 12
Q96T70 PRELIMINARY; PRT; 102 AA.
AC Q96T70
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Endostatin variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Deininger M.H., Trautmann K., Schluesser H.J.;
RT "Endostatin promotes delayed secondary damage following traumatic
RT brain injury."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF333247; AAK50626.1; -.
FT NON_TER
SQ SEQUENCE 102 AA; 11147 MW; ECAC47AA6420947D CRC64;
Query Match 59.2%; Score 529; DB 4; Length 102;
Best Local Similarity 98.0%; Pred. No. 2.2e-46;
Matches 100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 59 IVNLKDELLFPSEWALFSGSEGLKPGARIFSGDKVLRHPTPQKSVWHGSDPNRRRL 118

```



DR SMART; SMO0210; TSPN; 1.  
 KW Colligden.  
 SQ SEQUENCE 1388 AA; 141757 MW; 96828E45E847194B CRC64;  
 Query Match 56.6%; Score 505; DB 4; Length 1388;  
 Best Local Similarity 56.9%; Pred. No. 1.7e-42;  
 Matches 95; Conservative 28; Mismatches 40; Indels 4; Gaps 1;  
 QY 2 ALNSPLSGMGRGNGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIVN 61  
 Db 1222 ALNMPVSGDIR---ADFCQFQARAAGLLSTYRAFLSSHLQDLSTIVKAEYSLPIVN 1277  
 QY 62 LKDELLPFSWEALFSGSEGLKPGARIFSEGDGKDVLRHPTWPKSVWHGSDPNRRILTES 121  
 Db 1278 LKGOVLNNWDSIFSGHGGQFNHPIYSEFGRDIMTDPSPQKVIWHGSSPHGVRIVDN 1337  
 QY 122 YCETWRTAPSATCQASLLGGRLLGOSAAASHAYIVLICIENSFMT 168  
 Db 1338 YCEAWRTADTAVTGLASPLSTGKILDOKAYSCANRLIVLICIENSFMT 1384

Search completed: March 13, 2004, 08:17:54  
 Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 08:27:58 ; Search time 58 Seconds  
(without alignments)  
828.156 Million cell updates/sec

Title: US-09-171-607A-1

Perfect score: 893

Sequence: 1 VALNSPLSGMRGIRGADFQ.....ASCHYAVIVLCIENSFWTAS 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	893	100.0	178	3	AAU94324 Alternate
2	893	100.0	178	4	AAU00900 Human End
3	893	100.0	179	4	AAU00901 Human End
4	893	100.0	182	3	AAU59622 Human end
5	893	100.0	182	3	AAU59622 Human end
6	893	100.0	182	3	AAU59622 Human end
7	893	100.0	182	3	AAU59622 Human end
8	893	100.0	182	3	AAU59622 Human end
9	893	100.0	182	3	AAU59622 Human end
10	893	100.0	183	2	AAU02113 SEQ ID 76
11	893	100.0	183	2	AAU08693 Human end
12	893	100.0	183	3	AAU70252 Human ang
13	893	100.0	183	3	AAU90771 Human ang
14	893	100.0	183	3	AAU16451 Human end
15	893	100.0	183	3	AAU30493 Amino aci
16	893	100.0	183	4	AAU49379 Human end
17	893	100.0	183	4	AAU00896 Human End
18	893	100.0	183	5	AAU79901 Human end
19	893	100.0	183	5	AAU49503 Human end
20	893	100.0	183	5	AAU48995 Human end
21	893	100.0	183	5	AAU97132 Human end
22	893	100.0	183	6	AAU79753 Human end
23	893	100.0	193	3	AAU76690 Synthetic
24	893	100.0	216	3	AAU90874 Human HMW
25	893	100.0	275	5	AAU76689 Amino aci
					AAU76689 Synthetic

## ALIGNMENTS

RESULT 1

AAU94324

ID AAY94324 standard; protein; 178 AA.

XX AC AAY94324;

XX DT 11-AUG-2000 (first entry)

XX DE Alternate human endostatin protein.

XX KW Human; endothelial cell proliferation inhibitor; collagen XVIII;

XX KW angiogenesis inhibitor; anti-tumour; cytostatic; antiproliferative;

XX KW vasotropic; dermatological; ophthalmological; vulnerary;

XX KW antiarteriosclerotic; antidiabetic; haemostatic; contraceptive;

XX KW ocular angiogenic disease; atherosclerosis; scleroderma;

XX KW myocardial angiogenesis; telangiectasia; angiofibroma; wound granulation.

XX OS Homo sapiens.

XX PN WO200026368-A2.

XX PD 11-MAY-2000.

XX PF 01-NOV-1999; 99WO-US025605.

XX PR 30-OCT-1998; 98US-0106343P.

XX PR 20-MAY-1999; 99US-00315689.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI O'reilly MS, Folkman MJ;

XX DR WPI; 2000-365617/31.

XX DR N-FSDB; AAA27005.

XX PT Novel endostatin capable of inhibiting endothelial cell proliferation and

XX PT angiogenesis, useful for treating angiogenesis-dependent cancers and as

XX PT birth control agents.

XX PS Claim 3; Page 39; 68pp; English.

XX CC The present sequence is an alternate functional endostatin protein. When

XX CC the human endostatin gene sequence AAA27004 is recombinantly expressed,

XX CC an observable doublet of protein results, both versions of which are

XX CC functional endostatin proteins. The present endostatin variant is the

XX CC same as the protein encoded by AAA27004 minus the first four amino acids.

XX CC Recombinant mouse endostatin (20 mg/kg) was administered subcutaneously

XX CC to mice implanted with Lewis lung carcinomas. There was tumour mass

AAU76688 Human col  
ABG73586 Human End  
ABP41878 Human ova  
AAW26327 Human alp  
AAY25113 Human alp  
AAO17357 Human alp  
AAW92296 Human col  
AAY08694 Human col  
ABP96308 Human col  
ABP83471 Human col  
ABP68617 Human pan  
AAU00898 Human End  
AAU49810 Human End  
ABG31794 Human end  
AAG78717 Human vas  
AAU00899 Human End  
AAW90877 Human HMW  
AAB08407 A human a  
ADE40256 Human NOV  
AAY70265 Canine an

26 893 100.0 310 5 AAU76688  
27 893 100.0 513 5 ABG73586  
28 893 100.0 682 2 ABP41878  
29 893 100.0 684 2 AAW26327  
30 893 100.0 684 2 AAY25113  
31 893 100.0 684 5 AAO17357  
32 893 100.0 1301 2 AAU92296  
33 893 100.0 1336 2 AAY08694  
34 893 100.0 1336 6 ABP96308  
35 893 100.0 1516 5 ABP83471  
36 893 100.0 1516 5 ABP68617  
37 889 99.6 181 4 AAU00898  
38 889 99.6 183 4 AAB49810  
39 889 99.6 184 5 ABG31794  
40 885 99.2 184 4 AAG78717  
41 885 99.1 180 4 AAU00899  
42 885 99.1 193 3 AAU90877  
43 882 98.8 271 3 AAB08407  
44 876 98.1 184 7 ADE40256  
45 783 87.7 184 3 AAY70265

CC regression non-detectable levels after 12 days of therapy due to the  
 CC angiogenesis inhibitory activity of endostatin. Thus the protein is  
 CC useful for treatment of angiogenesis- dependent cancers. The  
 CC polynucleotide and polypeptide sequences of this endostatin are useful  
 CC for treating and diagnosis of tumours, ocular angiogenic diseases, Osler-  
 CC Webber syndrome, myocardial angiogenesis, plaque neovascularisation,  
 CC telangiectasia, haemophilic joints, angiofibroma and wound granulation,  
 CC for treatment of diseases related to excessive or abnormal stimulation of  
 CC endothelial cells e.g. intestinal adhesions, atherosclerosis,  
 CC scleroderma. The protein may also be useful as a birth control agent by  
 CC reducing or preventing uterine vascularisation. The gene for endostatin  
 CC may be isolated from cells or tissue that express high levels of  
 CC endostatin, eg. tumour cells, by generating cDNA from mRNA using reverse  
 CC transcriptase and then amplifying the DNA sequence  
 XX  
 SQ Sequence 178 AA;

Query Match 100.0%; Score 893; DB 3; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRAFLSRLQDLYSIVRRADRAAIVPIV 60  
 DB 9 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRAFLSRLQDLYSIVRRADRAAIVPIV 68  
 QY 61 NLKDELLFPPSWALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 69 NLKDELLFPPSWALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 128  
 QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 DB 129 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 178

RESULT 2  
 AAU00900  
 ID AAU00900 standard; protein; 178 AA.  
 XX  
 AC AAU00900;  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human Endostatin(TM) N-terminal deletion mutant protein#2.  
 XX  
 KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;  
 KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;  
 KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;  
 KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;  
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;  
 KW retinopathy of prematurity; macular corneal graft rejection;  
 KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;  
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;  
 KW haemophilic joint; angiofibroma; wound granulation; variant; mutant;  
 KW mutin.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200119989-A2.  
 XX  
 XX 22-MAR-2001.  
 XX  
 XX 14-SEP-2000; 2000WO-US025166.  
 XX  
 XX 14-SEP-1999; 99US-0153698P.  
 XX  
 XX (ENTR-) ENTREMED INC.  
 XX  
 XX Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;  
 PI Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;  
 XX WPI; 2001-244802/25.  
 DR N-PSDB; AAS00868.  
 XX

PT Producing Endostatin protein for treating angiogenesis mediated diseases  
 PT such as solid tumors, comprises recombinantly producing the protein using  
 PT an expression system, and recovering and purifying the protein.  
 XX  
 PS Claim 5; Page 33; 67pp; English.

XX The sequence represents Human Endostatin(TM) N-terminal deletion mutant  
 CC protein lacking the N-terminal 4 amino acids and the C-terminal lysine, a  
 CC natural variant recovered from fermentations of Pichia pastoris cultures  
 CC harbouring an expression plasmid containing the Endostatin(TM) DNA  
 CC sequence given in AAS00868. The new method of the invention is useful for  
 CC producing, recovering and purifying Endostatin(TM) from biological  
 CC sources, such as biological fluids, tissues, cells, culture media, and  
 CC fermentation media. Endostatin(TM) is useful for treating angiogenesis  
 CC mediated diseases such as solid tumours, blood borne tumours, leukaemias,  
 CC tumour metastases, benign tumours, e.g. haemangioma, acoustic arthritis,  
 CC neurofibromas, trachoma, and pyogenic granulomas, rheumatoid arthritis,  
 CC psoriasis, ocular angiogenic diseases, e.g., diabetic retinopathy,  
 CC retinopathy of prematurity, macular degeneration, corneal graft  
 CC rejection, neovascular glaucoma, colon cancer, retrolental fibroplasia,  
 CC rubeosis, Osler-Webber Syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
 CC and wound granulation. Endostatin(TM) is also useful for treating disease  
 CC of excessive or abnormal stimulation of endothelial cells such as  
 CC intestinal adhesions, atherosclerosis, scleroderma and hypertrophic  
 CC scars. Higher yields of more purified, and biologically active  
 CC Endostatin(TM) are obtained by the new method. Endostatin(TM) can be  
 CC stored in buffers for extended periods of time, and also subjected to  
 CC lyophilisation, while preserving biological activity. Centrifugation of  
 CC broth from fermentation steps in production is avoided, preventing  
 CC unwanted potential cellular lysis and contamination with additional  
 CC proteins, pigments, enzymes and other cellular chemicals and debris  
 XX  
 SQ Sequence 178 AA;

Query Match 100.0%; Score 893; DB 4; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRAFLSRLQDLYSIVRRADRAAIVPIV 60  
 DB 9 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRAFLSRLQDLYSIVRRADRAAIVPIV 68  
 QY 61 NLKDELLFPPSWALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 69 NLKDELLFPPSWALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 128  
 QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 DB 129 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 178

RESULT 3  
 AAU00901  
 ID AAU00901 standard; protein; 179 AA.  
 XX  
 AC AAU00901;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human Endostatin(TM) N-terminal mutant protein#1.  
 XX  
 KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;  
 KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;  
 KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;  
 KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;  
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;  
 KW retinopathy of prematurity; macular corneal graft rejection;  
 KW neovascular glaucoma; retrolental fibroplasia; Osler-Webber Syndrome;  
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;  
 KW haemophilic joint; angiofibroma; wound granulation; mutant; mutin.  
 XX  
 OS Homo sapiens.

XX WO200119989-A2.  
 XX  
 XX  
 PD 22-MAR-2001.  
 XX  
 XX 14-SEP-2000; 2000WO-US025166.  
 XX  
 XX 14-SEP-1999; 99US-0153698P.  
 XX  
 XX (ENTR-) ENTREMED INC.  
 XX  
 XX Liang H, Sim KL, Chang-Murad A, Zhou X, Madsen J, Boerner RJ;  
 PI Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;  
 XX  
 DR WPI: 2001-244802/25.  
 DR N-PSDB; AAS00868.  
 XX  
 XX  
 PT Producing Endostatin protein for treating angiogenesis mediated diseases  
 PT such as solid tumors, comprises recombinantly producing the protein using  
 PT an expression system, and recovering and purifying the protein.  
 XX  
 PS Claim 5; Page 32; 67pp; English.  
 XX  
 CC The sequence represents a Human Endostatin(TM) N-terminal deletion mutant  
 CC lacking the N-terminal 4 amino acids. The new method of the invention is  
 CC useful for producing, recovering and purifying Endostatin (TM) from  
 CC biological sources, such as biological fluids, tissues, cells, culture  
 CC media, and fermentation media. Endostatin(TM) is useful for treating  
 CC angiogenesis mediated diseases such as solid tumors, blood borne  
 CC tumours, leukaemias, tumour metastases, benign tumours, e.g. haemangioma,  
 CC acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas,  
 CC rheumatoid arthritis, psoriasis, ocular angiogenic diseases, e.g.,  
 CC diabetic retinopathy, retinopathy of prematurity, macular degeneration,  
 CC corneal graft rejection, neovascular glaucoma, colon cancer, retrolental  
 CC fibroplasia, rubecosis, Osler-Webber Syndrome, myocardial angiogenesis,  
 CC plaque neovascularisation, telangiectasia, haemophilic joints,  
 CC angiofibroma, and wound granulation. Endostatin(TM) is also useful for  
 CC treating disease of excessive or abnormal stimulation of endothelial  
 CC cells such as intestinal adhesions, atherosclerosis, scleroderma and  
 CC hypertrophic scars. Higher yields of more purified, and biologically  
 CC active Endostatin(TM) are obtained by the new method. Endostatin(TM) can  
 CC be stored in buffers for extended periods of time, and also subjected to  
 CC lyophilisation, while preserving biological activity. Centrifugation of  
 CC broth from fermentation steps in production is avoided, preventing  
 CC unwanted potential cellular lysis and contamination with additional  
 CC proteins, pigments, enzymes and other cellular chemicals and debris  
 XX  
 SQ Sequence 179 AA;  
 Query Match 100.0%; Score 893; DB 4; Length 179;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGADFCQFQOARAVGLAGTFRFLSSRLQDLYSVRRADRAAIVP 60  
 DB 9 VALNSPLSGMGRGADFCQFQOARAVGLAGTFRFLSSRLQDLYSVRRADRAAIVP 68  
 QY 61 NLKDELLFPSWEALFSGSEGLPKPGARIFSDGKDLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 69 NLKDELLFPSWEALFSGSEGLPKPGARIFSDGKDLRHPTWPKSVWHGSDPNRRLTE 128  
 QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWMTAS 170  
 DB 129 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWMTAS 178  
 RESULT 4  
 AAY59622  
 ID AAY59622 standard; protein; 182 AA.  
 XX  
 AC AAY59622;  
 XX  
 DT 14-MAR-2000 (first entry)

XX Human endostatin protein fragment.  
 DE  
 XX  
 KW Endostatin; scatter factor activity; human; tubulogenesis; psoriasis;  
 KW metastatic cancer; tumorigenesis; ocular angiogenic disease;  
 KW rheumatoid arthritis; Osler-Webber syndrome; telangiectasia;  
 KW haemophilic joint; angiofibroma; wound granulation.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9962944-A2.  
 XX  
 XX 09-DEC-1999.  
 XX  
 XX 03-JUN-1999; 99WO-US012278.  
 XX  
 PR 03-JUN-1998; 98US-0087890P.  
 PR 10-JUL-1998; 98US-0092393P.  
 PR 01-SEP-1998; 98US-0098790P.  
 XX  
 PA (CHIL-) CHILDRENS MEDICAL CENT.  
 XX  
 XX Javaherian K, Folkman MJ;  
 PI WPI: 2000-072833/06.  
 XX  
 DR New endostatin oligomers, used for treating e.g. tumors.  
 PT  
 PT  
 PS Disclosure; Page 6; 44pp; English.  
 XX  
 CC This sequence is a fragment of the human endostatin protein. Endostatin  
 CC is an approximately 20kD C-terminal globular domain of the collagen-like  
 CC protein collagen XVII. Protein oligomers consisting of more than one  
 CC endostatin monomer have anti-tubulogenic effects and induce  
 CC reorganization of the actin cytoskeleton. The oligomer has scatter factor  
 CC activity. The oligomers induce the destruction of tubular lumens and  
 CC elongation of cells, and inhibit tubulogenesis and tumorigenesis. The  
 CC oligomers can also be used to treat metastatic cancers, tumours,  
 CC rheumatoid arthritis, psoriasis, ocular angiogenic disease, Osler-Webber  
 CC syndrome, plaque neovascularisation, telangiectasia, haemophilic joints,  
 CC angiofibroma and wound granulation. The oligomers can also be used to  
 CC treat diseases that have angiogenesis as a pathological consequence e.g.  
 CC ulcers. The endostatin oligomers can also be used to develop affinity  
 CC columns for isolating antibodies or receptors. Passive antibody therapy  
 CC using antibodies that specifically bind endostatin oligomers can be used  
 CC to modulate morphogenic processes such as metastatic cancer as well as  
 CC angiogenesis-dependent processes such as reproduction, development, wound  
 CC healing, tissue repair, and angiogenesis-dependent diseases. Also,  
 CC antisera directed to the Fab regions of endostatin oligomer antibodies  
 CC can be administered to block the ability of endogenous endostatin  
 CC oligomer antisera to bind endostatin oligomers  
 XX  
 SQ Sequence 182 AA;  
 Query Match 100.0%; Score 893; DB 3; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGADFCQFQOARAVGLAGTFRFLSSRLQDLYSVRRADRAAIVP 60  
 DB 13 VALNSPLSGMGRGADFCQFQOARAVGLAGTFRFLSSRLQDLYSVRRADRAAIVP 72  
 QY 61 NLKDELLFPSWEALFSGSEGLPKPGARIFSDGKDLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 73 NLKDELLFPSWEALFSGSEGLPKPGARIFSDGKDLRHPTWPKSVWHGSDPNRRLTE 132  
 QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWMTAS 170  
 DB 133 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWMTAS 182  
 RESULT 5  
 AAY94323



ID AAY94323 standard; protein; 182 AA.  
 AC AAY94323;  
 XX  
 XX  
 DT 11-AUG-2000 (first entry)  
 XX  
 DE Human endostatin protein.  
 XX  
 KW Human; endothelial cell proliferation inhibitor; collagen XVIII;  
 KW angiogenesis inhibitor; anti-tumour; cytostatic; antiproliferative;  
 KW vasotrophic; dermatological; ophthalmological; vulnerary;  
 KW antiarteriosclerotic; antidiabetic; haemostatic; contraceptive;  
 KW ocular angiogenic disease; atherosclerosis; scleroderma;  
 KW myocardial angiogenesis; telangiectasia; angiofibroma; wound granulation.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200026368-A2.  
 FN  
 PD 11-MAY-2000.  
 XX  
 PF 01-NOV-1999; 99WO-US025605.  
 XX  
 PR 30-OCT-1998; 98US-0106343P.  
 PR 20-MAY-1999; 99US-00315689.  
 XX  
 XX (CHIL-) CHILDRENS MEDICAL CENT.  
 PA  
 XX O'reilly MS, Folkman MJ;  
 PI  
 XX MPI; 2000-365617/31.  
 DR N-PSDB; AAA27004.  
 XX  
 XX Novel endostatin capable of inhibiting endothelial cell proliferation and  
 PT angiogenesis, useful for treating angiogenesis-dependent cancers and as  
 PT birth control agents.  
 XX  
 PS Claim 2; Page 38; 68pp; English.  
 XX  
 CC The present sequence is an endostatin protein which is the carboxy  
 CC terminal protein of human collagen XVIII. Recombinant mouse endostatin  
 CC (20 mg/kg) was administered subcutaneously to mice implanted with Lewis  
 CC lung carcinomas. There was tumour mass regression non-detectable levels  
 CC after 12 days of therapy due to the angiogenesis inhibitory activity of  
 CC endostatin. Thus the protein is useful for treatment of angiogenesis-  
 CC dependent cancers. The polynucleotide and polypeptide sequences of this  
 CC endostatin are useful for treating and diagnosis of tumours, ocular  
 CC angiogenic diseases, Osler-Webber syndrome, myocardial angiogenesis,  
 CC plaque neovascularisation, telangiectasia, haemophilic joints,  
 CC angiofibroma and wound granulation, for treatment of diseases related to  
 CC excessive or abnormal stimulation of endothelial cells e.g. intestinal  
 CC adhesions, atherosclerosis, scleroderma. The protein may also be useful  
 CC as a birth control agent by reducing or preventing uterine  
 CC vascularisation. The gene for endostatin may be isolated from cells or  
 CC tissue that express high levels of endostatin, eg. tumour cells, by  
 CC generating cDNA from mRNA using reverse transcriptase and then amplifying  
 CC the DNA sequence  
 XX  
 SQ Sequence 182 AA;  
 Query Match 100.0%; Score 893; DB 3; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCFQQAARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60  
 Db 13 VALNSPLSGMGRGIRGADFCFQQAARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72  
 QY 61 NLKDELLFPSSWEALFSGSEGPLKPGARIFSPGKQVLRHPTWPKQSVWHGSDPNGRRLTE 120  
 Db 73 NLKDELLFPSSWEALFSGSEGPLKPGARIFSPGKQVLRHPTWPKQSVWHGSDPNGRRLTE 132  
 QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWMTAS 170  
 Db 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWMTAS 182  
 RESULT 7  
 AAU00897  
 ID AAU00897 standard; protein; 182 AA.  
 XX  
 AC AAU00897;

Db 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWMTAS 182  
 RESULT 6  
 AAB28399  
 ID AAB28399 standard; protein; 182 AA.  
 XX  
 AC AAB28399;  
 XX  
 DT 19-FEB-2001 (first entry)  
 XX  
 DE Human endostatin.  
 XX  
 KW Human; endostatin; cytostatic; antiproliferative;  
 KW vascular endothelial growth factor; VEGF; antibody; VEGF2 receptor;  
 KW cancer; vascularised solid tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200064946-A2.  
 FN  
 PD 02-NOV-2000.  
 XX  
 PF 28-APR-2000; 2000WO-US011367.  
 PR  
 XX 28-APR-1999; 99US-0131432P.  
 XX  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA  
 XX Thorpe PE, Brekken RA;  
 PI  
 XX MPI; 2000-687317/67.  
 DR  
 XX Immunogenic composition for the treatment and diagnosis of cancer  
 PT comprises an anti-VEGF (vascular endothelial growth factor) antibody  
 PT binding the same epitope as the monoclonal antibody ATCC PTA 1595.  
 XX  
 PS Example 10; Page 291-292; 298pp; English.  
 XX  
 CC The present invention relates to anti-Vascular Endothelial Growth Factor  
 CC (VEGF) antibodies that bind to the same epitope as the monoclonal  
 CC antibody ATCC PTA 1595 and which significantly inhibit VEGF binding to  
 CC the VEGF receptor VEGFR2, without inhibiting VEGF binding to the VEGF  
 CC receptor VEGFR1. The present sequence is human endostatin. Endostatin may  
 CC be conjugated onto the anti-VEGF antibodies of the present invention. The  
 CC anti-VEGF antibodies of the present invention are useful for the  
 CC treatment and diagnosis of cancer, especially vascularised solid tumours  
 XX  
 SQ Sequence 182 AA;  
 Query Match 100.0%; Score 893; DB 3; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCFQQAARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 60  
 Db 13 VALNSPLSGMGRGIRGADFCFQQAARAVGLAGTFRAFLSSRLQDLYSIVRRADRAAVPIV 72  
 QY 61 NLKDELLFPSSWEALFSGSEGPLKPGARIFSPGKQVLRHPTWPKQSVWHGSDPNGRRLTE 120  
 Db 73 NLKDELLFPSSWEALFSGSEGPLKPGARIFSPGKQVLRHPTWPKQSVWHGSDPNGRRLTE 132  
 QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWMTAS 170  
 Db 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWMTAS 182  
 RESULT 7  
 AAU00897  
 ID AAU00897 standard; protein; 182 AA.  
 XX  
 AC AAU00897;

XX DT 04-JUL-2001 (first entry)  
 XX DE Human Endostatin(TM) C-terminus minus 1 protein.  
 XX  
 KW Human; Endostatin(TM); angiogenesis mediated disease; solid tumours;  
 KW blood borne tumour; leukaemia; tumour metastasis; benign tumour;  
 KW haemangioma; acoustic neuroma; neurofibroma; trachoma; rubeosis;  
 KW pyogenic granuloma; rheumatoid arthritis; psoriasis; colon cancer;  
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;  
 KW retinopathy of prematurity; macular corneal graft rejection;  
 KW neovascular glaucoma; retrolental fibroplasia; Osler-Weber Syndrome;  
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;  
 KW haemophilic joint; angiofibroma; wound granulation; variant;  
 KW C-terminus minus 1 protein.  
 XX OS Homo sapiens.  
 XX PN WO200119989-A2.  
 XX PD 22-MAR-2001.  
 XX PF 14-SEP-2000; 2000WO-US025166.  
 XX PR 14-SEP-1999; 99US-0153698P.  
 XX PA (ENTR-) ENTREMED INC.  
 XX PI Liang H, Sim KL, Chang-Murad A, Zhou X, Madson J, Boerner RJ;  
 XX PI Bermejo LL, Mistry FR, Shepard SR, Schrimsher JL;  
 XX DR WPI: 2001-244802/25.  
 XX DR N-PSDB; AA500897.  
 XX  
 XX PT Producing Endostatin protein for treating angiogenesis mediated diseases  
 XX PT such as solid tumors, comprises recombinantly producing the protein using  
 XX PT an expression system, and recovering and purifying the protein.  
 XX  
 XX PS Claim 5; Page 30; 67pp; English.  
 XX  
 CC The sequence represents Human Endostatin(TM) C-terminus minus 1 protein,  
 CC a natural variant lacking the C-terminal amino acid of Endostatin(TM)  
 CC recovered from fermentations of Pichia pastoris cultures harbouring a  
 CC expression plasmid containing the Endostatin(TM) DNA sequence given in  
 CC AA800867. The new method of the invention is useful for producing,  
 CC recovering and purifying Endostatin(TM) from biological sources, such as  
 CC biological fluids, tissues, cells, culture media, and fermentation media.  
 CC Endostatin(TM) is useful for treating angiogenesis mediated diseases such  
 CC as solid tumours, blood borne tumours, leukaemias, tumour metastases,  
 CC benign tumours, e.g. haemangioma, acoustic neuromas, neurofibromas,  
 CC trachomas, and pyogenic granulomas, rheumatoid arthritis, psoriasis,  
 CC ocular angiogenic diseases, e.g., diabetic retinopathy, retinopathy of  
 CC prematurity, macular degeneration, corneal graft rejection, neovascular  
 CC glaucoma, colon cancer, retrolental fibroplasia, rubeosis, Osler-Weber  
 CC Syndrome, myocardial angiogenesis, plaque neovascularisation,  
 CC telangiectasia, haemophilic joints, angiofibroma, and wound granulation.  
 CC Endostatin(TM) is also useful for treating disease of excessive or  
 CC abnormal stimulation of endothelial cells such as intestinal adhesions,  
 CC atherosclerosis, scleroderma and hypertrophic scars. Higher yields of  
 CC more purified, and biologically active Endostatin(TM) are obtained by the  
 CC new method. Endostatin(TM) can be stored in buffers for extended periods  
 CC of time, and also subjected to lyophilisation, while preserving  
 CC biological activity. Centrifugation of broth from fermentation steps in  
 CC production is avoided, preventing unwanted potential cellular lysis and  
 CC contamination with additional proteins, pigments, enzymes and other  
 CC cellular chemicals and debris  
 XX  
 XX Sequence 182 AA;

Query Match 100.0%; Score 893; DB 4; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLACTFRFLSRLQDLYSIVRRADRAAVPIV 60  
 |||||  
 DB 13 VALNSPLSGMGRGIRGADFCQFOQARAVGLACTFRFLSRLQDLYSIVRRADRAAVPIV 72  
 |||||  
 QY 61 NLKDELLFSPSEALFSGSGPLKPGARIFSFQDKVLRHPTWPKSVWHGSDPNGRRLTE 120  
 |||||  
 DB 73 NLKDELLFSPSEALFSGSGPLKPGARIFSFQDKVLRHPTWPKSVWHGSDPNGRRLTE 132  
 |||||  
 QY 121 SYCETWRTAPSATGQASLLGRLGQSAASHHAYIVLCIENSFWMTAS 170  
 |||||  
 DB 133 SYCETWRTAPSATGQASLLGRLGQSAASHHAYIVLCIENSFWMTAS 182  
 |||||  
 RESULT 8  
 AAU77951  
 ID AAU77951 standard; protein; 182 AA.  
 XX  
 AC AAU77951;  
 XX  
 DT 02-JUL-2002 (first entry)  
 XX  
 DE Amino acid sequence for human endostatin.  
 XX  
 KW Human; immunoconjugate; anti-vascular endothelial growth factor antibody;  
 KW anti-VEGF antibody; monoclonal antibody 2C3 ATCC PTA 1595; VEGF receptor;  
 KW VEGFR2; KDR/Flk-1; VEGFR1; Flt-1; angiogenesis; macular degeneration;  
 KW ocular neovascular disease; cancer; vascularised solid tumour; AIDS;  
 KW metastatic tumour; endothelial cell proliferation; inflammatory disorder;  
 KW atherosclerosis; diabetic retinopathy; corneal graft rejection;  
 KW acquired immune deficiency syndrome; infection; restenosis; fungal ulcer;  
 KW sickle cell anaemia; endometriosis; endostatin.  
 KW  
 OS Homo sapiens.  
 XX  
 PN AU200179401-A.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 12-OCT-2001; 2001AU-00079401.  
 XX  
 PR 28-APR-2000; 2000AU-00048049.  
 XX  
 PA (TEXA) UNIV TEXAS SYSTEM.  
 XX  
 PI Thorpe PE, Brekken RA;  
 XX  
 DR WPI: 2002-281368/33.  
 XX  
 PT Immunoconjugate compositions for treating cancer by inhibiting  
 PT angiogenesis and for delivering a diagnostic agent to tumor, comprises  
 PT anti-vascular endothelial growth factor antibody attached to a biological  
 PT agent.  
 XX  
 XX Example 10; Page 12-13 (Sequence listing); 299pp; English.  
 XX  
 CC The present invention relates to antibody-based compositions comprising  
 CC an immunoconjugate such as anti-vascular endothelial growth factor (VEGF)  
 CC antibody (Ab) (or its antigen-binding fragment), attached to a biological  
 CC agent, where the Ab binds to the same epitope as the monoclonal antibody  
 CC (MAb) 2C3 ATCC PTA 1595, and significantly inhibits VEGF binding to the  
 CC VEGF receptor VEGFR2 (KDR/Flk-1) without inhibiting VEGF binding to the  
 CC VEGF receptor VEGFR1 (Flt-1). The compositions of the invention are  
 CC useful in therapy, and diagnosis, for inhibiting angiogenesis in an  
 CC animal having ocular neovascular disease or macular degeneration, and for  
 CC delivering a biological agent to a vascularised tumour. The compositions  
 CC can also be used for treating cancer and subjects at risk of developing,  
 CC a vascularised solid tumour, a metastatic tumour or metastases from a  
 CC VEGF-induced endothelial cell proliferation, without significantly inhibiting  
 CC VEGF-induced macrophage, osteoclast or chondroblast function.  
 CC The compositions can be used for treating various diseases such as  
 CC inflammatory disorders, atherosclerosis, diabetic retinopathy,  
 CC restenosis, acquired immune deficiency syndrome (AIDS), blood borne

CC tumors, corneal graft rejection, Crohn's disease, fungal ulcers,  
 CC infections, sickle cell anemia, and endometriosis. The present sequence  
 CC represents human endostatin. Endostatin may be attached or functionally  
 CC associated with anti-VEGF antibodies  
 XX  
 SQ Sequence 182 AA;

Query Match 100.0%; Score 893; DB 5; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSLRLQDLYSIVRRADAAVPIV 60  
 Db 13 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSLRLQDLYSIVRRADAAVPIV 72

QY 61 NLKDELLFPSWEALFSGSEGLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRLTE 120  
 Db 73 NLKDELLFPSWEALFSGSEGLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRLTE 132

QY 121 SYCETWTEAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSPMTAS 170  
 Db 133 SYCETWTEAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSPMTAS 182

RESULT 9  
 AAY02113  
 ID AAY02113 standard; protein; 183 AA.  
 XX AC AAY02113;  
 XX DT 16-JUL-1999 (first entry)  
 XX DE SEQ ID 76 of WO9916889.  
 XX KW Angiostatin; endostatin; interferon; thrombospondin;  
 KW interferon-inducible protein; platelet factor 4; anti-angiogenic;  
 KW anti-tumor; multifunctional protein; angiogenic-mediated disease; cancer;  
 KW diabetic retinopathy; macular degeneration; arthritis;  
 KW tumor cell production.  
 XX OS Homo sapiens.  
 XX PN WO9916889-A1.  
 XX PD 08-APR-1999.  
 XX PF 30-SEP-1998; 98WO-US020464.  
 XX PR 01-OCT-1997; 97US-0060609P.  
 XX PA (SEAR ) SEARLE & CO G D.  
 XX PI Bolanowski MA, Caparon MH, Casperson GF, Gregory SA, Klein BK;  
 PI McKearn JP;  
 XX WPI; 1999-255098/21.  
 XX PT New multifunctional proteins useful for treating angiogenic-mediated  
 PT diseases.  
 XX PS Disclosure; Page 106-107; 121pp; English.  
 XX CC The specification describes multifunctional proteins which comprise  
 CC combinations of angiostatin, endostatin, interferon, thrombospondin,  
 CC interferon-inducible protein and platelet factor 4, and have anti-  
 CC angiogenic and/or anti-tumor activity. The multifunctional protein may  
 CC exhibit useful properties such as having similar or greater biological  
 CC activity when compared to a single factor or by having improved half-life  
 CC or decreased adverse side effects, or a combination of these properties.  
 CC The proteins can be used for treating an angiogenic-mediated disease,  
 CC e.g. cancer, diabetic retinopathy, macular degeneration, or arthritis.  
 CC They can also be used for inhibiting the production of tumor cells  
 CC (characteristic of lung, breast, ovarian, prostate, pancreatic, gastric,

CC colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma)  
 CC in a patient and for inhibiting tumor growth. The present sequence is  
 CC used in the course of the invention  
 XX  
 SQ Sequence 183 AA;

Query Match 100.0%; Score 893; DB 2; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSLRLQDLYSIVRRADAAVPIV 60  
 Db 13 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSLRLQDLYSIVRRADAAVPIV 72

QY 61 NLKDELLFPSWEALFSGSEGLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRLTE 120  
 Db 73 NLKDELLFPSWEALFSGSEGLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRLTE 132

QY 121 SYCETWTEAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSPMTAS 170  
 Db 133 SYCETWTEAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSPMTAS 182

RESULT 10  
 AAY08693  
 ID AAY08693 standard; protein; 183 AA.  
 XX AC AAY08693;  
 XX DT 10-AUG-1999 (first entry)  
 XX DE Human endostatin protein fragment.  
 XX KW Plasminogen; human; angiotensin; endostatin; gene therapy; vector;  
 KW anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;  
 KW tumor growth; solid tumour; diabetic retinopathy; retina.  
 XX OS Homo sapiens.  
 XX PN WO9926480-A1.  
 XX PD 03-JUN-1999.  
 XX PF 20-NOV-1998; 98WO-US024950.  
 XX PR 20-NOV-1997; 97US-00975424.  
 XX PA (GENE-) GENETIX PHARM INC.  
 XX PA (MASI ) MASSACHUSETTS INST TECHNOLOGY.  
 XX PI Leboulch P, Pawliuk RJ, Bachelot T;  
 XX WPI; 1999-357696/30.  
 XX DR N-PSDB; AAX7719.  
 XX PT Anti-angiogenic gene therapy vectors.  
 XX PS Disclosure; Page 74-75; 83pp; English.  
 XX CC This invention describes a novel viral gene therapy vector comprising a  
 CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen from  
 CC human or murine angiostatin, human or murine endostatin and angiogenesis-  
 CC inhibiting fusions and fragments, where the viral vector is sufficiently  
 CC attenuated for use in human gene therapy. The products of the invention  
 CC have anti-angiogenic, cytostatic, anti-diabetic and ophthalmological  
 CC activity. The vector is used in gene therapy for inhibiting tumour growth  
 CC in humans harbouring a solid tumour. The vector expresses an anti-  
 CC angiogenic polypeptide. An additional use comprises treatment of diabetic  
 CC retinopathy, where the anti-angiogenic polypeptide inhibits angiogenesis  
 CC in the vicinity of the retina. The vector is administered to cells ex  
 CC vivo and then administered to the patient  
 XX  
 SQ Sequence 183 AA;

Query Match 100.0%; Score 893; DB 2; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMRGIRGADFCQFQOQARAVGLAGTFRFLSSRLQDLYSIIVRRADRAAVPIV 60  
 DB 13 VALNSPLSGMRGIRGADFCQFQOQARAVGLAGTFRFLSSRLQDLYSIIVRRADRAAVPIV 72

QY 61 NLKDELLFFSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
 DB 73 NLKDELLFFSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASHHAYIVLCIENSFWMTAS 170  
 DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASHHAYIVLCIENSFWMTAS 182

RESULT 11  
 AAY70252  
 ID AAY70252 standard; protein; 183 AA.  
 XX AC AAY70252;  
 XX DT 06-JUN-2000 (first entry)  
 XX DE Human angiogenesis inhibitor, endostatin.  
 XX KW Human; immunoglobulin gamma Fc fragment; endostatin; immunofusin;  
 XX KW angiogenesis; inhibitor; cytosolic; antirheumatoid; antiarthritic;  
 XX KW antipsoriasis; antidiabetic; ophthalmological; immunosuppressant;  
 XX KW vasotropic; vulnerary; treatment; antiarteriosclerosis; tumour;  
 XX KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;  
 XX KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;  
 XX KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;  
 XX KW wound granulation; keloid scar; gene therapy.  
 XX OS Homo sapiens.  
 XX PN WO200011033-A2.  
 XX PD 02-MAR-2000.  
 XX PF 25-AUG-1999; 99WO-US019329.  
 XX PR 25-AUG-1998; 98US-0097883F.  
 XX PA (LEXI-) LEXINGEN PHARM CORP.  
 XX PI Lo K, Li Y, Gillies SD;  
 XX DR WPI; 2000-237616/20.  
 XX DR N-PSDB; AAZ51291.

Novel fusion protein of angiotensin or endostatin and an immunoglobulin FC region, useful for treating conditions mediated by angiogenesis, such as rheumatoid arthritis, tumors and macular degeneration.

Example 1; Page 41-42; 68pp; English.

The patent discloses a DNA molecule encoding a fusion protein comprising a signal sequence, an immunoglobulin Fc region, and an angiogenesis inhibitor selected from angiotensin, endostatin, a plasminogen fragment having angiotensin activity, a collagen XVIII fragment having endostatin activity, or combinations of them. The fusion protein (immunofusin) is used to inhibit angiogenesis and to treat diseases or conditions mediated by angiogenesis. Conditions that may be treated include solid tumors, blood born tumors, tumor metastasis, benign tumors including haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases e.g. diabetic retinopathy, retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis and Osler-Webber syndrome; myocardial angiogenesis,

CC plaque neovascularisation, telangiectasia, haemophilic joints;  
 CC angiofibroma, wound granulation, and excessive or abnormal stimulation of  
 CC endothelial cells, intestinal cells, atherosclerosis, sclerodermal and  
 CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used in  
 CC gene therapy. The present sequence is a human endostatin used in the  
 CC construction of immunofusin containing human immunoglobulin gamma (IgG)  
 CC Fc fragment  
 XX SQ Sequence 183 AA;

Query Match 100.0%; Score 893; DB 3; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMRGIRGADFCQFQOQARAVGLAGTFRFLSSRLQDLYSIIVRRADRAAVPIV 60  
 DB 13 VALNSPLSGMRGIRGADFCQFQOQARAVGLAGTFRFLSSRLQDLYSIIVRRADRAAVPIV 72

QY 61 NLKDELLFFSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
 DB 73 NLKDELLFFSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASHHAYIVLCIENSFWMTAS 170  
 DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASHHAYIVLCIENSFWMTAS 182

RESULT 12  
 AAY90771  
 ID AAY90771 standard; protein; 183 AA.  
 XX AC AAY90771;  
 XX DT 22-AUG-2000 (first entry)  
 XX DE Human angiogenesis inhibiting factor 1 protein.  
 XX KW Human; angiogenesis inhibiting factor 1; IAF-1; tumour; antibody;  
 XX KW abnormal vessel disease.  
 XX OS Homo sapiens.  
 XX PN CN1244536-A.  
 XX PD 16-FEB-2000.  
 XX PF 10-AUG-1998; 98CN-00117150.  
 XX PR 10-AUG-1998; 98CN-00117150.  
 XX PA (ONCO-) INST ONCOLOGY UNDER TUMOR HOSPITAL CHINE.  
 XX PI Yang Z, Guo W;  
 XX DR WPI; 2000-388168/34.  
 XX DR N-PSDB; AAA29884.

Angiogenesis inhibiting factor 1 and its derivative useful for treating tumors.

Claim 1; Fig 5; 41pp; Chinese.

The present sequence represents an angiogenesis inhibiting factor (I), designated IAF-1. The present invention also describes: (1) preparation of (I) and its derivative; (2) an IAF binding acceptor and its preparation; and (3) an IAF antibody. (I) is useful for preparing new biological preparations for effectively treating various tumours and abnormal-vessel diseases. The IAF antibody is preferably a polyclonal antibody, mosaic antibody, single stranded antibody and human originated antibody

Sequence 183 AA;

Query Match 100.0%; Score 893; DB 3; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADAAVPIV 60  
 DB 13 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADAAVPIV 72

QY 61 NLKDELLFSPWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 73 NLKDELLFSPWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHGSDPNRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWTAS 170  
 DB 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWTAS 182

RESULT 14  
 AAB16451  
 ID AAB16451 standard; protein; 183 AA.  
 AC AAB16451;  
 DT 27-OCT-2000 (first entry)  
 XX Human endostatin protein sequence.

XX Angiogenesis-inhibiting protein receptor; angiogenesis; angiostatin;  
 endostatin; plasminogen; laminin; treatment; wound healing; solid tumour;  
 psoriasis; scleroderma; myocardial angiogenesis; Crohn's disease;  
 cerebral collateral; arteriovenous malformation; rubecosis; cancer;  
 diabetic retinopathy; arthritis; wound healing; peptic ulcer;  
 Helicobacter related disease; fracture; cat scratch fever.

XX Homo sapiens.  
 XX WO200032631-A2.  
 XX 08-JUN-2000.  
 XX 06-DEC-1999; 99WO-US028897.  
 XX 04-DEC-1998; 98US-00206059.  
 XX (ENTR-) ENTREMED INC.  
 XX Macdonald NJ, Sim KL;  
 XX WPI; 2000-412290/35.

XX New angiogenesis-inhibiting protein receptors, useful in methods for  
 treating diseases and processes that are mediated by angiogenesis, such  
 as solid tumors, psoriasis, scleroderma and myocardial angiogenesis.

XX Disclosure; Fig 3; 100pp; English.

XX This invention relates to angiogenesis-inhibiting protein receptors, and  
 the DNA sequences encoding them. Angiogenesis is the generation of new  
 blood vessels into a tissue, and normally occurs in wound healing, foetal  
 and embryonal development and the formation of the corpus luteum,  
 endometrium and placenta. Angiostatin is a protein (see AAB16450 and  
 AAA68202) involved in angiogenesis, and has an amino acid sequence  
 similar to that of a plasminogen fragment (see murine plasminogen  
 AAB16490). Angiostatin has the ability to inhibit angiogenesis.  
 XX Endostatin is also an angiogenesis inhibiting protein (see AAB16451 and  
 AAA68203). Sequences AAB68242 and AAB16522 represent coding and protein  
 sequences of human laminin. Laminin is an angiostatin binding protein,  
 and some of the peptides of the invention share homology with regions of  
 laminin. Peptides AAB16452-B16521 (excluding AAB16490) are the  
 XX angiogenesis-inhibiting protein receptor fragments of the invention. The  
 peptides bind either angiostatin or endostatin and can be used in methods  
 for treating diseases and processes that are mediated by angiogenesis,  
 XX such as solid tumors, psoriasis, scleroderma, myocardial angiogenesis,

CC Crohn's disease, cerebral collaterals, arteriovenous malformations,  
 CC rubecosis, diabetic retinopathy, arthritis, wound healing, peptic ulcers,  
 CC Helicobacter related diseases, fractures, placentation and cat scratch  
 CC fever. They are useful for the detection and prognosis of cancer. DNA  
 CC sequences A628204-A628241 encode the peptides of the invention  
 XX Sequence 183 AA;

Query Match 100.0%; Score 893; DB 3; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADAAVPIV 60  
 DB 13 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADAAVPIV 72

QY 61 NLKDELLFSPWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 73 NLKDELLFSPWEALFSGSEGLPKPGARIFSPDGKDVLRHPTWPKSVWHGSDPNRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWTAS 170  
 DB 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWTAS 182

RESULT 14  
 AAB30493  
 ID AAB30493 standard; protein; 183 AA.  
 AC AAB30493;  
 DT 06-MAR-2001 (first entry)  
 XX Amino acid sequence of human endostatin encoded by plasmid pMALch#15.  
 XX Streptomyces sp. strain C5; SnpA; S. venezuelae; alpha-amylase;  
 endostatin; cancer; tumour growth; angiogenesis.  
 XX Homo sapiens.  
 XX WO2000060945-A1.  
 XX 19-OCT-2000.  
 XX 12-APR-2000; 2000WO-US009747.  
 XX 13-APR-1999; 99US-0129084P.  
 XX (MERI) MERCK & CO INC.  
 XX Desanti CL, Strohl WR;  
 XX WPI; 2000-686970/67.  
 XX N-PSDB; AAC62023.

XX Preparation of soluble recombinant endostatin involves transforming  
 Streptomyces host with expression vector comprising nucleotide sequence  
 encoding endostatin operably linked to linker and leader peptide.

XX Example 1; Fig 6; 57pp; English.

XX The present sequence represents human endostatin. The protein is  
 expressed in Streptomyces. Leader sequences of Streptomyces sp. strain C5  
 SnpA and S. venezuelae alpha-amylase proteins are linked to the N-  
 terminal of endostatin. This ensures that endostatin protein is produced  
 as a secreted, soluble protein which needs no refolding, is stable in the  
 fermentation broth and is produced in large quantities. The method is  
 used for preparing soluble recombinant human, murine or primate  
 endostatin, which is useful in the treatment of cancer, inhibition of  
 tumour growth, inhibition of angiogenesis, isolation of receptors for  
 endostatin and for identification of anti-angiogenic compounds in assays.  
 XX The endostatin protein is produced as a secreted, soluble protein which  
 needs no refolding, is stable in the fermentation broth and is produced



cancer, retrolental fibroplasia, rubeosis, Osler-Webber Syndrome.  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilic joints, angiodiroma, and wound granulation. Endostatin(TM)  
 CC is also useful for treating disease of excessive or abnormal stimulation  
 CC of endothelial cells such as intestinal adhesions, atherosclerosis,  
 CC scleroderma and hypertrophic scars. Higher yields of more purified, and  
 CC biologically active Endostatin(TM) are obtained by the new method.  
 CC Endostatin(TM) can be stored in buffers for extended periods of time, and  
 CC also subjected to lyophilisation, while preserving biological activity.  
 CC Centrifugation of broth from fermentation steps in production is avoided,  
 CC preventing unwanted potential cellular lysis and contamination with  
 CC additional proteins, pigments, enzymes and other cellular chemicals and  
 CC debris  
 CC  
 XX Sequence 183 AA;  
 SQ  
 Query Match 100.0%; Score 893; DB 4; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60  
 DB 13 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 72  
 QY 61 NLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 73 NLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVLRHPTWPKSVWHGSDPNRRLTE 132  
 QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVILCIENSFMTAS 170  
 DB 133 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVILCIENSFMTAS 182  
 RESULT 17  
 ABB79901  
 ID ABB79901 standard; protein; 183 AA.  
 XX  
 AC ABB79901;  
 XX  
 DT 05-DEC-2002 (first entry)  
 XX  
 DE Human endostatin polypeptide.  
 XX  
 KW Endostatin; human; ophthalmological; ocular neovascularisation;  
 KW choroidal neovascularisation; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200267971-A2.  
 XX  
 PD 06-SEP-2002.  
 XX  
 PF 21-FEB-2002; 2002WO-US0053336.  
 XX  
 PR 22-FEB-2001; 2001US-0270787P.  
 PR 04-APR-2001; 2001US-0281296P.  
 XX  
 XX (NOVS ) NOVARTIS AG.  
 PA  
 XX  
 XX Brazzell RK, Campochiario PA, Dixon KH;  
 PI  
 XX  
 DR WPI; 2002-698636/75.  
 DR N-PSDB; ABQ81193.  
 XX  
 XX Treating or preventing choroidal neovascularization comprises increasing  
 PT the amount of endostatin in ocular tissues of afflicted individuals to a  
 PT choroidal neovascularization inhibiting level.  
 XX  
 XX Claim 2; Page 39; 44pp; English.  
 PS  
 XX The present sequence is the protein sequence of a human endostatin  
 CC polypeptide. A claimed method for the treatment of ocular  
 CC neovascularisation, especially choroidal neovascularisation, involves

increasing the level of endostatin in ocular tissue, especially where the  
 endostatin comprises the present sequence, or is its fragment, derivative  
 or variant. The increase is effected by administering a viral vector,  
 especially an adenovirus, adeno-associated virus, a retrovirus or  
 lentivirus vector, comprising an endostatin-encoding nucleic acid. Cells  
 secreting endostatin may be encapsulated and implanted within an  
 individual. The method is used when ocular neovascularisation is caused  
 by histoplasmosis, pathological myopia, angiod streaks, anterior  
 ischaemic optic neuropathy, bacterial endocarditis, Best's disease,  
 birdshot retinochoroidopathy, choroidal haemangioma, choroidal naevi,  
 choroidal nonperfusion, choroidal osteomas, choroidal rupture,  
 chorioideraemia, chronic retinal detachment, coloboma of the retina,  
 Drusen, endogenous Candida endophthalmitis, extrapillary hamartoma of  
 the retinal pigmented epithelium, fundus flavimaculatus, idiopathic,  
 macular hole, malignant melanoma, membranoproliferative glomerulonephritis  
 (type II), metallic intraocular foreign body, morning glory disc  
 syndrome, multiple evanescent white-dot syndrome, neovascularisation of  
 ora serrata, operating microscope burn, optic nerve head pits,  
 photocoagulation, punctate inner choroidopathy, rubella, sarcoidosis,  
 serpinginous or geographic choroiditis, subretinal fluid drainage, tiled  
 disc syndrome, Toxoplasma retinochoroiditis, tuberculosis, Vogt-Koyanagi-  
 Harada syndrome, diabetic retinopathy, non-diabetic retinopathy, brain  
 vein occlusion, central retinal vein occlusion, retinopathy in premature  
 infants, rubeosis iridis, neovascular glaucoma, perifoveal  
 telangiectasis, sickle cell retinopathy, Sallé's disease, retinal  
 vasculitis, Von Hippel Lindau disease, radiation retinopathy, retinal  
 cryol injury, retinitis pigmentosa, retinochoroidal coloboma, corneal  
 neovascularisation due to herpes simplex keratitis, corneal ulcers,  
 keratoplasty, pterygia and trauma (all claimed)  
 CC  
 XX Sequence 183 AA;  
 SQ  
 Query Match 100.0%; Score 893; DB 5; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60  
 DB 13 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 72  
 QY 61 NLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 73 NLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVLRHPTWPKSVWHGSDPNRRLTE 132  
 QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVILCIENSFMTAS 170  
 DB 133 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVILCIENSFMTAS 182  
 RESULT 18  
 AAM49503  
 ID AAM49503 standard; protein; 183 AA.  
 XX  
 AC AAM49503;  
 XX  
 DT 07-MAY-2002 (first entry)  
 XX  
 DE Human endostatin protein.  
 XX  
 KW Endostatin; human; proliferation; blood vessel endothelium; regeneration;  
 KW tumour; blood vessel; treatment; amplification.  
 XX  
 OS Homo sapiens.  
 XX  
 PN CN1177005-A.  
 XX  
 PD 25-MAR-1998.  
 XX  
 PF 10-SEP-1997; 97CN-00107112.  
 PR 10-SEP-1997; 97CN-00107112.  
 XX  
 XX (XUGG/) XU G.



XX Xu G, Ren M, Xu L;  
 XX WPI; 2002-106746/15.  
 DR N-FSDB; ABA99261.  
 XX  
 PT Gene clone of inhibitory factor for hyperplasia of inner blood vessel  
 PT cells in human body's real tumor, and its use in anti-tumor blood vessel  
 PT regeneration.  
 XX  
 XX Disclosure; Page 4 (Disclosure); 6pp; Chinese.  
 XX  
 CC This invention describes a novel preparation which inhibits the  
 CC proliferation of blood vessel endothelium and prevents the regeneration  
 CC activity of tumor blood vessels. The preparation can also be used as a  
 CC biological preparation in the treatment of tumors. This sequence  
 CC represents the human endostatin protein described in the invention  
 XX  
 SQ Sequence 183 AA;  
 Query Match 100.0%; Score 893; DB 5; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 VALNSPLSGMGRGIRGADFCFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIV 60  
 Db 13 VALNSPLSGMGRGIRGADFCFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIV 72  
 Qy 61 NLKDELLFPSSWEALFSGSEGPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNGRRLTE 120  
 Db 73 NLKDELLFPSSWEALFSGSEGPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNGRRLTE 132  
 Qy 121 SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 Db 133 SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSFMTAS 182  
 RESULT 19  
 AAM48895  
 ID AAM48895 standard; protein; 183 AA.  
 XX  
 AC AAM48895;  
 XX  
 DT 04-APR-2002 (first entry)  
 XX  
 DE Human endostatin protein.  
 XX  
 KW Human; angiostatin; endostatin; angiogenesis; cancer; metastasis;  
 KW psoriasis; scleroderma; Crohn's disease; corneal disease; retinopathy;  
 KW arthritis; wound healing; Helicobacter pylori; peptic ulcer;  
 KW gene therapy; angiostatin antagonist; endostatin antagonist;  
 KW antiangiogenic; cytostatic; antiarthritis; antiinflammatory;  
 KW cerebroprotective; antidiabetic; virucide; antipyretic; vulnerary;  
 KW gynaecological; cat scratch fever.  
 XX  
 OS Homo sapiens.  
 XX  
 FT WO200193897-A2.  
 XX  
 FT 13-DEC-2001.  
 XX  
 PD 04-JUN-2001; 2001WO-US017947.  
 XX  
 PF 02-JUN-2000; 2000US-0209065P.  
 XX  
 PR 08-MAY-2001; 2001US-0289387P.  
 XX  
 XX (ENTR-) ENTREMED INC.  
 XX  
 PA Sim KL, Macdonald NJ;  
 XX  
 PI Sim KL, Liang H;  
 XX  
 DR WPI; 2002-130569/17.  
 XX  
 PT Regulating angiogenesis and treatment of angiogenesis-mediated diseases,

PT e-g. hemangioma, tumors or cancer, by administering a tropomyosin binding  
 PT compound or actin disrupting compound.  
 XX  
 XX Disclosure; Fig 3; 95pp; English.  
 XX  
 CC The present invention relates to methods of regulating angiogenesis in an  
 CC individual by administering an angiogenesis regulating composition  
 CC comprising a tropomyosin binding compound or an actin disrupting  
 CC compound. The compositions are useful for treating diseases and processes  
 CC mediated by angiogenesis including haemangioma, solid tumours, blood  
 CC bourne tumours, leukaemia, metastasis, Crohn's disease, coronary or  
 CC cerebral collaterals, arthritis, diabetic neovascularisation, macular  
 CC degeneration, wound healing, Helicobacter related diseases, ovulation,  
 CC menstruation, and cat scratch fever. The present sequence is a protein  
 CC described in the exemplification of the invention  
 XX  
 SQ Sequence 183 AA;  
 Query Match 100.0%; Score 893; DB 5; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 VALNSPLSGMGRGIRGADFCFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIV 60  
 Db 13 VALNSPLSGMGRGIRGADFCFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIV 72  
 Qy 61 NLKDELLFPSSWEALFSGSEGPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNGRRLTE 120  
 Db 73 NLKDELLFPSSWEALFSGSEGPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNGRRLTE 132  
 Qy 121 SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 Db 133 SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAYIVLCIENSFMTAS 182  
 RESULT 20  
 AAU97132  
 ID AAU97132 standard; protein; 183 AA.  
 XX  
 AC AAU97132;  
 XX  
 DT 13-AUG-2002 (first entry)  
 XX  
 DE Human endostatin.  
 XX  
 KW Human; angiogenesis; PITSLRE protein kinase; cancer; arthritis;  
 KW macular degeneration; diabetic retinopathy; angiogenic-related disease;  
 KW haemangioma; blood borne tumour; leukaemia; neovascularisation;  
 KW coronary collateral; cerebral collateral; neovascular glaucoma;  
 KW corneal disease; wound healing; Helicobacter related disease; fracture;  
 KW keloid; ovulation; menstruation.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 64..126  
 FT /note= "Internal peptide of endostatin is a PITSLRE  
 FT homologous region"  
 XX  
 XX WO200230982-A2.  
 XX  
 XX 18-APR-2002.  
 XX  
 XX 15-OCT-2001; 2001WO-US032437.  
 XX  
 XX 13-OCT-2000; 2000US-0240127P.  
 XX  
 XX (ENTR-) ENTREMED INC.  
 XX  
 XX Sim KL, Liang H;  
 XX  
 XX WPI; 2002-435440/46.  
 DR  
 DR N-FSDB; ABK50685.

```

XX PT Regulating angiogenesis for treating scleroderma, leukemia, keloids by
XX PT administering a protein that is homologous to PITSRE protein kinase and
XX PT an angiogenic factor or a protein kinase and its active fragments.
XX PS Disclosure; Fig 2A; 45pp; English.
XX CC The present invention relates to methods and compositions of inhibiting
XX CC angiogenesis. The method comprises administering to a human or animal a
XX CC composition comprising a protein that is homologous to PITSRE protein
XX CC kinases (PK) and an angiogenic factor. The method is useful for
XX CC regulating angiogenesis related to cancer, arthritis, macular
XX CC degeneration, and diabetic retinopathy. The compositions are useful for
XX CC inhibiting angiogenic-related diseases. The method and compositions are
XX CC useful in treating diseases and processes that are mediated by
XX CC angiogenesis including haemangioma, solid tumours, blood borne tumours,
XX CC leukaemia, metastasis, telangiectasia, psoriasis, scleroderma, pyogenic
XX CC granuloma, myocardial angiogenesis, Crohn's disease, plaque
XX CC neovascularisation, coronary collaterals, cerebral collaterals,
XX CC arteriovenous malformations, ischaemic limb angiogenesis, corneal
XX CC diseases, rubeosis, neovascular glaucoma, diabetic retinopathy,
XX CC retrolental fibroplasia, arthritis, diabetic neovascularisation, macular
XX CC degeneration, wound healing, peptic ulcer, Helicobacter related diseases,
XX CC fractures, keloids, vasculogenesis, haematopoiesis, ovulation,
XX CC menstruation, placentalation, and cat scratch fever. The method of the
XX CC invention provides a therapy for cancer that has minimal side effects.
XX CC The present sequence represents human endostatin which is used to
XX CC generate angiogenesis-inhibiting peptides
XX SQ Sequence 183 AA;

Query Match 100.0%; Score 893; DB 5; Length 183;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 60
DB 13 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 72

QY 61 NLKDELLFSEWALFSGSEGPLKPGARIFSPDGKDLRHPTWPKSVHMGSDPNGRRLTE 120
DB 73 NLKDELLFSEWALFSGSEGPLKPGARIFSPDGKDLRHPTWPKSVHMGSDPNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASHHAYIVILCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASHHAYIVILCIENSFMTAS 182

RESULT 21
AAG79753
ID AAG79753 standard; protein; 183 AA.
XX AC AAG79753;
XX DT 18-MAR-2003 (first entry)
XX DE Human endostatin.
XX KW Human; plasminogen; angiostatin; neovascularisation; kringle domain;
XX KW cell proliferation; viral vector; replication-defective; cancer; tumour.
XX OS Homo sapiens.
XX PN WO200288173-A2.
XX PD 07-NOV-2002.
XX PF 29-APR-2002; 2002WO-US013461.
XX PR 30-APR-2001; 2001US-0287673P.
XX PR 05-APR-2002; 2002US-0370634P.
XX PA (CELL-) CELL GENESYS INC.

XX PI Chang B, Wu WW, Macarthur J, Patel S, Jooss K, Mendez M;
XX WPI; 2003-129131/12.
XX DR N-PSDB; ABA00774.
XX PT New recombinant viral vector expressing human angiostatin useful for
XX PT inhibiting angiogenesis in a mammalian subject with cancer or tumor.
XX PS Example 4; Page 82-83; 83pp; English.
XX CC This sequence represents endostatin. Endostatin is a 20 kD C-terminal
XX CC fragment of collagen XVIII that inhibits angiogenesis. The endostatin
XX CC coding sequence may be used in the recombinant viral vector of the
XX CC invention for obtaining angiostatin activity. The vector comprises a
XX CC promoter capable of expressing human angiostatin operably linked to a
XX CC structural gene encoding one or more domains of human angiostatin. The
XX CC vector, which may be a replication-defective viral vector, is useful for
XX CC inhibiting angiogenesis in a mammal, especially with cancer or a tumour
XX SQ Sequence 183 AA;

Query Match 100.0%; Score 893; DB 6; Length 183;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 60
DB 13 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 72

QY 61 NLKDELLFSEWALFSGSEGPLKPGARIFSPDGKDLRHPTWPKSVHMGSDPNGRRLTE 120
DB 73 NLKDELLFSEWALFSGSEGPLKPGARIFSPDGKDLRHPTWPKSVHMGSDPNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASHHAYIVILCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASHHAYIVILCIENSFMTAS 182

RESULT 22
AAU76690
ID AAU76690 standard; protein; 193 AA.
XX AC AAU76690;
XX DT 21-MAY-2002 (first entry)
XX DE Synthetic plasmid pEnd-HR#2 FPD fusion protein sequence.
XX KW Mouse; Ig signal peptide; mIgSP; functional protein domain; FPD;
XX KW primary translational product; PTP; DNA construct; regulatory DNA;
XX KW DNA targeting segment; regulatory factor; single regulatory unit;
XX KW monoclonal antibody; recombination-derived alteration; blood product;
XX KW human; COL18A1; mutant; mutein; fusion protein.
XX OS Mus sp.
XX OS Homo sapiens.
XX OS Synthetic.
XX OS Chimeric.
XX FH Key Location/Qualifiers
XX FT Peptide 1..19 /label= Signal peptide
XX FT FT /note= "Mouse Ig signal peptide (mIgSP)"
XX FT Region 1..19 /note= "Encoded by mouse Ig signal peptide (mIgSP) exon"
XX FT Protein 20..193 /label= Mature human COL18A1 protein"
XX FT FT /note= "Contains exons 39-41"
XX FT Region 20..84 /note= "Encoded by human COL18A1 exon 39"
XX FT Region 85..124 /note= "Encoded by human COL18A1 exon 40"

```

FT Region 125. .193  
 XX /note= "Encoded by human COL18A1 exon 41"

PN WO200210372-A1.

PD 07-FEB-2002.

PF 01-AUG-2001; 2001WO-GB003455.

PR 01-AUG-2000; 2000GB-00018876.

XX (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.  
 PA (CHAP/) CHAPMAN P W.

PI Chapman PW, De Luca G, Falciola L;

XX MPI; 2002-195963/25.

DR N-PSDB; ABK09978.

XX Producing functional protein domain by growing host cell transfected with  
 PT DNA construct having regulatory DNA and DNA targeting segment, and  
 PT optionally culturing homologically recombinant cell and collecting  
 PT protein.

XX Example; Fig 9; 116pp; English.

XX The present invention relates to a new method of producing a protein,  
 CC such as functional protein domain, that is either C- or N-terminus of the  
 CC primary translational product (PTP) of a gene, where the protein has  
 CC biological activity which is distinct from PTP. The method of the  
 CC invention involves growing a host cell transfected with a DNA construct  
 CC comprising a regulatory DNA and a DNA targeting segment. This method is  
 CC useful for producing a functional protein domain of proteins such as  
 CC regulatory factors, blood products and monoclonal antibodies. The method  
 CC described in the invention allows controlled and precise modification of  
 CC the host cell genome in order to produce functional protein domain (FPD).  
 CC The amount of exogenous sequence to be integrated in the host cell genome  
 CC is very limited since, as coding sequence, the original coding sequence  
 CC present in the host cell genome itself is used. Use of the host cell  
 CC sequence encoding FPD also provides the advantages of both eliminating  
 CC any recombination-derived alteration of such coding sequence, and also  
 CC making use of the same post-transcriptional (e.g., splicing) and/or post-  
 CC translational (e.g. glycosylation, phosphorylation) processes that are  
 CC actually applied in vivo for the maturation of FPD. The use of a single  
 CC regulatory unit eliminates the necessity of manipulating the  
 CC complementary DNA coding for the PTP to isolate the segment coding for  
 CC the FPD, and adapt it to the expression vector. The present amino acid  
 CC sequence represents the plasmid pmd-HR#2 FPD fusion protein of the  
 CC invention. This fusion protein contains the mouse Ig signal peptide  
 CC (migSP) sequence fused to exons 39-41 of the human COL18A1 sequence

XX Sequence 193 AA;

Query Match 100.0%; Score 893; DB 5; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGADFCQFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIV 60  
 DB 23 VALNSPLSGMGRIGADFCQFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIV 82  
 QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSPGKDVLRHPTWPKSVHMGSDPNRRRLTE 120  
 DB 83 NLKDELLFPSWEALFSGSEGPKPGARIFSPGKDVLRHPTWPKSVHMGSDPNRRRLTE 142  
 QY 121 SYCETWRTAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
 DB 143 SYCETWRTAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 192

RESULT 23

AAW90874

ID AAW90874 standard; protein; 195 AA.

XX AAW90874;  
 XX 07-JUL-2000 (first entry)  
 XX Human HMW endostatin (1) protein.  
 XX Endostatin; human; renal insufficiency; antitumor; antiproliferative;  
 KW treatment; angiogenesis; tumor; vascular disease.  
 XX Homo sapiens.  
 XX WO200017240-A1.  
 XX 30-MAR-2000.  
 XX 21-SEP-1999; 99WO-EP006963.  
 XX 21-SEP-1998; 98DE-01042992.  
 PR 03-APR-1999; 99DE-01015267.  
 PR 08-JUN-1999; 99DE-01026040.  
 XX (HAEM-) HAEMOPEP PHARMA GMBH.  
 XX Staendker L, Forssmann W;  
 XX MPI; 2000-292826/25.  
 XX New high molecular weight form of endostatin, useful e.g. as  
 PT antiangiogenic agent for treating cancer, isolated from hemofiltrate of  
 PT patients with kidney failure.  
 XX Claim 2; Page 18-19; 32pp; German.

XX This invention describes novel human high molecular weight (HMW)  
 CC endostatin (HE) protein fragments isolated from the hemofiltrate of  
 CC patients with renal insufficiency. The products of the invention have  
 CC antitumor and antiproliferative activity. HE is used to treat; (i)  
 CC diseases that involve uncontrolled angiogenesis, particularly tumors; and  
 CC (ii) vascular diseases of supporting or connective tissue, respiratory  
 CC tract, cardiovascular system, urogenital tract and nervous system, or  
 CC sensory organs (particularly the eye). HE is also used to raise specific  
 CC antibodies which are used for diagnosis and treatment of conditions that  
 CC involve overexpression of HE. HE has a very long plasma half-life and can  
 CC be administered repeatedly without inducing an immune response. AAW90874-  
 CC W90877 represent the endostatin proteins described in the method of the  
 CC invention

XX Sequence 195 AA;

Query Match 100.0%; Score 893; DB 3; Length 195;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRIGADFCQFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIV 60  
 DB 26 VALNSPLSGMGRIGADFCQFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIV 85  
 QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSPGKDVLRHPTWPKSVHMGSDPNRRRLTE 120  
 DB 86 NLKDELLFPSWEALFSGSEGPKPGARIFSPGKDVLRHPTWPKSVHMGSDPNRRRLTE 145  
 QY 121 SYCETWRTAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
 DB 146 SYCETWRTAPSATQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 195

RESULT 24

AAB30495

ID AAB30495 standard; protein; 216 AA.

XX AAB30495;

XX

DT 06-MAR-2001 (first entry)  
 XX Amino acid sequence of vaa-endostatin fusion protein in pANT3052.  
 DE Streptomyces sp. strain C5; SnpA; S. venezuelae; alpha-amylase;  
 XX endostatin; cancer; tumour growth; angiogenesis.  
 KW Synthetic.  
 KW Homo sapiens.  
 OS Streptomyces sp.  
 OS Homo sapiens.  
 XX Key Location/Qualifiers  
 PH Peptide 1..28  
 FT /note= "vaa signal sequence"  
 FT Protein 29..216  
 FT /note= "endostatin"  
 XX WO200060945-A1.  
 XX 9-OCT-2000.  
 XX 12-APR-2000; 2000WO-US009747.  
 XX 13-APR-1999; 99US-0129084P.  
 XX (MERI ) MERCK & CO INC.  
 XX Desanti CL, Strohl WR;  
 PI WPI: 2000-686970/67.  
 DR N-PSDB; AAC62025.  
 XX Preparation of soluble recombinant endostatin involves transforming  
 PT Streptomyces host with expression vector comprising nucleotide sequence  
 PT encoding endostatin operably linked to linker and leader peptide.  
 XX Example 1; Fig 10A-B; 57pp; English.  
 XX The present sequence represents a fusion protein of vaa and endostatin.  
 CC The specification describes a method for the production of soluble,  
 CC recombinant human endostatin in Streptomyces. Leader sequences of  
 CC Streptomyces sp. strain C5 SnpA and S. venezuelae alpha-amylase proteins  
 CC are linked to the N-terminal of endostatin. This ensures that endostatin  
 CC protein is produced as a secreted, soluble protein which needs no  
 CC refolding, is stable in the fermentation broth and is produced in large  
 CC quantities. The method is used for preparing soluble recombinant human,  
 CC murine or primate endostatin, which is useful in the treatment of cancer,  
 CC inhibition of tumour growth, inhibition of angiogenesis, isolation of  
 CC receptors for endostatin and for identification of anti-angiogenic  
 CC compounds in assays. The endostatin protein is produced as a secreted,  
 CC soluble protein which needs no refolding, is stable in the fermentation  
 CC broth and is produced in large quantities. Streptomyces are amenable  
 CC for cultivation in large fermentations allowing for large quantities of  
 CC soluble endostatin to be produced  
 XX Sequence 216 AA;  
 SQ Query Match 100.0%; Score 893; DB 3; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-101;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSRLQDLYSIVRRADRAAVPIV 60  
 DB 46 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRFLSRLQDLYSIVRRADRAAVPIV 105  
 QY 61 NLKDELLFPSEWALFSGSEGLPKPGARIFSGDKGVLRHPTWPKSWHGSQDPNGRLTE 120  
 DB 106 NLKDELLFPSEWALFSGSEGLPKPGARIFSGDKGVLRHPTWPKSWHGSQDPNGRLTE 165  
 QY 121 SYCETWRTAPSGATGQASSLLGRLGQSAASCHHAYIVLVCIENTSPWTAS 170  
 DB 166 SYCETWRTAPSGATGQASSLLGRLGQSAASCHHAYIVLVCIENTSPWTAS 215

RESULT 25  
 AAU76689  
 ID AAU76689 standard; protein; 275 AA.  
 XX AAU76689;  
 AC AAU76689;  
 XX 21-MAY-2002 (first entry)  
 DT Synthetic plasmid pEnd-HR#1 FPD fusion protein sequence.  
 DE Mouse; Ig signal peptide; mIgSP; functional protein domain; FPD;  
 XX Primary translational product; PTP; DNA construct; regulatory DNA;  
 KW DNA targeting segment; regulatory factor; single regulatory unit;  
 KW monoclonal antibody; recombination-derived alteration; blood product;  
 KW human; COL18A1; mutant; mutein; fusion protein.  
 XX Mus sp.  
 OS Homo sapiens.  
 OS Synthetic.  
 OS Chimeric.  
 XX Key Location/Qualifiers  
 PH Peptide 1..119  
 FT /label= Signal\_peptide  
 FT /note= "Mouse Ig signal peptide (mIgSP)"  
 FT Region 1..119  
 FT Protein 20..275  
 FT /label= Mature\_human\_COL18A1\_protein"  
 FT /note= "Contains exons 38-41"  
 XX WO200210372-A1.  
 PN 07-FEB-2002.  
 XX 01-AUG-2001; 2001WO-GB003455.  
 XX 01-AUG-2000; 2000GB-00018876.  
 XX (ISIT ) ARS APPLIED RES SYSTEMS HOLDING NV.  
 PA (CHAP/) CHAPMAN P W.  
 XX Chapman PW, De Luca G, Falcicola L;  
 PI WPI: 2002-195963/25.  
 XX N-PSDB; ABK09977.  
 DR Producing functional protein domain by growing host cell transfected with  
 XX DNA construct having regulatory DNA and DNA targeting segment, and  
 PT optionally culturing homologously recombinant cell and collecting  
 PT protein.  
 XX Example; Fig 8; 116pp; English.  
 PS The present invention relates to a new method of producing a protein,  
 XX such as functional protein domain, that is either C- or N-terminus of the  
 CC primary translational product (PTP) of a gene, where the protein has  
 CC biological activity which is distinct from PTP. The method of the  
 CC invention involves growing a host cell transfected with a DNA construct  
 CC comprising a regulatory DNA and a DNA targeting segment. This method is  
 CC useful for producing a functional protein domain of proteins such as  
 CC regulatory factors, blood products and monoclonal antibodies. The method  
 CC described in the invention allows controlled and precise modification of  
 CC the host cell genome in order to produce functional protein domain (FPD).  
 CC The amount of exogenous sequence to be integrated in the host cell genome  
 CC is very limited since, as coding sequence, the original coding sequence  
 CC present in the host cell genome itself is used. Use of the host cell  
 CC sequence encoding FPD also provides the advantages of both eliminating  
 CC any recombination-derived alteration of such coding sequence, and also  
 CC making use of the same post-transcriptional (e.g., splicing) and/or post-  
 CC translational (e.g. glycosylation, phosphorylation) processes that are  
 CC actually applied in vivo for the maturation of FPD. The use of a single

CC regulatory unit eliminates the necessity of manipulating the  
 CC complementary DNA coding for the PTP to isolate the segment coding for  
 CC the FPD, and adapt it to the expression vector. The present amino acid  
 CC sequence represents the plasmid pBnd-HR#1 FPD fusion protein of the  
 CC invention. This fusion protein contains the mouse Ig signal peptide  
 CC (mIgSp) sequence fused to exons 38-41 of the human COL18A1 sequence  
 XX Sequence 275 AA;

Query Match 100.0%; Score 893; DB 5; Length 275;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-101; Indels 0; Gaps 0;  
 Matches 170; Conservative 0; Mismatches 0;

QY 1 VALNSPLSGMGRGIRGADPQCQARAVGLAGTFRFLSSRLQDLYSIVRRADRAVPV 60  
 Db 105 VALNSPLSGMGRGIRGADPQCQARAVGLAGTFRFLSSRLQDLYSIVRRADRAVPV 164

QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSGDKVLRHPTWPKSVWHGSDPNGRRLTE 120  
 Db 165 NLKDELLFPSWEALFSGSEGPKPGARIFSGDKVLRHPTWPKSVWHGSDPNGRRLTE 224

QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
 Db 225 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 274

RESULT 26  
 AAU76688  
 ID AAU76688 standard; protein; 310 AA.  
 AC AAU76688;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE Human collagen XVIII alpha NCl domain protein sequence.  
 XX  
 KW Human; collagen XVIII alpha NCl domain; functional protein domain; FPD;  
 KW primary translational product; PTP; DNA construct; regulatory DNA;  
 KW DNA targeting segment; regulatory factor; single regulatory unit;  
 KW monoclonal antibody; recombination-derived alteration; blood product.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Domain 1..54 /label= Multimerisation\_domain  
 FT Region 1..43 /note= "Encoded by exon 36"  
 FT Region 44..54 /note= "Encoded by exon 37"  
 FT Region 55..136 /note= "Encoded by exon 38"  
 FT Region 137..310 /note= "Encoded by exon 39"  
 FT Region 137..310 /note= "Encoded by exon 40"  
 FT Region 137..310 /note= "Encoded by exon 41"  
 FT Region 137..201 /note= "Autonomous folding unit"  
 FT Region 202..241 /note= "Encoded by exon 39"  
 FT Region 242..310 /note= "Encoded by exon 40"  
 FT Region 242..310 /note= "Encoded by exon 41"

WO200210372-A1.  
 XX  
 PN  
 XX  
 PD 07-FEB-2002.  
 XX  
 XX  
 PF 01-AUG-2001; 2001WO-GB003455.  
 XX  
 PR 01-AUG-2000; 2000GB-00018876.  
 XX  
 XX (ISTE ) ARS APPLIED RES SYSTEMS HOLDING NV.  
 PA (CHAP/) CHAPMAN P W.

XX Chapman PW, De Luca G, Falciola L;  
 XX WPI; 2002-195963/25.  
 DR  
 XX  
 PT Producing functional protein domain by growing host cell transfected with  
 PT DNA construct having regulatory DNA and DNA targeting segment, and  
 PT optionally culturing homologously recombinant cell and collecting  
 PT protein.  
 XX  
 PS Example; Fig 3; 116pp; English.  
 XX  
 CC The present invention relates to a new method of producing a protein,  
 CC such as functional protein domain, that is either C- or N-terminus of the  
 CC primary translational product (PTP) of a gene, where the protein has  
 CC biological activity which is distinct from PTP. The method of the  
 CC invention involves growing a host cell transfected with a DNA construct  
 CC comprising a regulatory DNA and a DNA targeting segment. This method is  
 CC useful for producing a functional protein domain of proteins such as  
 CC regulatory factors, blood products and monoclonal antibodies. The method  
 CC described in the invention allows controlled and precise modification of  
 CC the host cell genome in order to produce functional protein domain (FPD).  
 CC The amount of exogenous sequence to be integrated in the host cell genome  
 CC is very limited since, as coding sequence, the original coding sequence  
 CC present in the host cell genome itself is used. Use of the host cell  
 CC sequence encoding FPD also provides the advantages of both eliminating  
 CC any recombination-derived alteration of such coding sequence, and also  
 CC making use of the same post-transcriptional (e.g., splicing) and/or post-  
 CC translational (e.g., glycosylation, phosphorylation) processes that are  
 CC actually applied in vivo for the maturation of FPD. The use of a single  
 CC regulatory unit eliminates the necessity of manipulating the  
 CC complementary DNA coding for the PTP to isolate the segment coding for  
 CC the FPD, and adapt it to the expression vector. The present amino acid  
 CC sequence represents the human collagen XVIII alpha NCl, a functional  
 CC protein domain used to illustrate the method of the invention  
 XX  
 SQ Sequence 310 AA;

Query Match 100.0%; Score 893; DB 5; Length 310;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-101; Indels 0; Gaps 0;  
 Matches 170; Conservative 0; Mismatches 0;

QY 1 VALNSPLSGMGRGIRGADPQCQARAVGLAGTFRFLSSRLQDLYSIVRRADRAVPV 60  
 Db 140 VALNSPLSGMGRGIRGADPQCQARAVGLAGTFRFLSSRLQDLYSIVRRADRAVPV 199

QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSGDKVLRHPTWPKSVWHGSDPNGRRLTE 120  
 Db 200 NLKDELLFPSWEALFSGSEGPKPGARIFSGDKVLRHPTWPKSVWHGSDPNGRRLTE 259

QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
 Db 260 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 309

RESULT 27  
 ABG73586  
 ID ABG73586 standard; protein; 513 AA.  
 XX  
 AC ABG73586;  
 XX  
 DT 03-MAR-2003 (first entry)  
 XX  
 DE Human Endostatin/IgG1Fc fusion construct.  
 XX  
 KW Human; endostatin; IgG1Fc; tumour; vascular endothelial proliferation;  
 KW vascular endothelial cytopoiesis inhibiting factor; inhibitor;  
 KW fusion construct.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers

FT Misc-difference 20 /note= "Encoded by CGF"  
 FT Misc-difference 271. .277  
 FT /note= "Encoded by TAGTAA"  
 XX  
 PN CN1354186-A.  
 XX  
 PD 19-JUN-2002.  
 XX  
 PF 30-NOV-2000; 2000CN-00123347.  
 XX  
 PR 30-NOV-2000; 2000CN-00123347.  
 XX  
 PA (LIAO-) LIAONING WEIXING BIOLOGICAL PROD INST CO.  
 XX  
 PI Chen L, Li Z, Liu Q;  
 XX  
 DR WPI; 2002-751441/82.  
 DR N-PSDB; ABQ76740.  
 XX  
 XX Preparation of recombinant human vascular endothelial cytopoiesis  
 PT suppressor factor with human IgG1Fc fragment molecular structure and  
 PT application of its product.  
 XX  
 PS Disclosure; Page 7-8 (Disclosure); 12pp; Chinese.  
 XX  
 CC This invention describes a novel method for the preparation of  
 CC recombinant human vascular endothelial cytopoiesis inhibiting factor with  
 CC human IgG1Fc fragment molecular structure and its product application.  
 CC The novel factor is derived from endostatin (using PCR to screen a human  
 CC foetal kidney cell cDNA library) and human IgG1Fc. The product of the  
 CC invention can specifically inhibit tumour vascular endothelial  
 CC proliferation and can be used for curing several tumour types. This  
 CC sequence represents a fusion construct composed of human endostatin and  
 CC human IgG1Fc, described in the disclosure of the invention  
 XX  
 SQ Sequence 513 AA;  
 Query Match 100.0%; Score 893; DB 5; Length 513;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCQFQQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60  
 DB 100 VALNSPLSGMGRGIRGADFCQFQQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 159  
 QY 61 NLKDELLFSPWEALFSGSGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
 DB 160 NLKDELLFSPWEALFSGSGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 219  
 QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 DB 220 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 269  
 RESULT 28  
 ABP41878  
 ID ABP41878 standard; protein; 682 AA.  
 XX  
 AC ABP41878;  
 XX  
 XX 22-AUG-2002 (first entry)  
 DT  
 XX Human ovarian antigen HEBBK29, SEQ ID NO:3010.  
 DE  
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;

KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive; chromosome 21q22.3.  
 OS Homo sapiens.  
 XX WO200200677-A1.  
 XX 03-JAN-2002.  
 XX 07-JUN-2001; 2001WO-US018569.  
 XX 07-JUN-2000; 2000US-0209467P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Birse CE, Rosen CA;  
 XX WPI; 2002-147878/19.  
 DR N-PSDB; ABQ54955.  
 XX  
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX  
 PS Claim 11; SEQ ID NO 3010; 2922pp; English.  
 XX  
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 682 AA;  
 Query Match 100.0%; Score 893; DB 5; Length 682;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-100;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCQFQQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60  
 DB 512 VALNSPLSGMGRGIRGADFCQFQQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 571  
 QY 61 NLKDELLFSPWEALFSGSGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
 DB 572 NLKDELLFSPWEALFSGSGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 631  
 QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 XX

Db 632 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 681

## RESULT 29

AAW26327  
ID AAW26327 standard; protein; 684 AA.

XX AC AAW26327;

XX DT 19-NOV-1997 (first entry)

XX DE Human alpha-1 collagen (XVIII).

XX KW Alpha-1 collagen; type XVIII collagen; cartilage degeneration.

XX OS Homo sapiens.

	Key	Location/Qualifiers
FT	Peptide	1..6
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	7..12
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	13..18
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	19..24
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	25..30
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	31..36
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	37..42
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	43..48
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	49..54
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	55..59
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	60..64
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	65..69
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	70..74
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	75..79
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	80..84
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	85..89
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	90..94
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	95..99
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	100..103
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	104..109
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	110..115
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	116..126
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	127..134
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	135..140
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif

FT	Peptide	/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	141..146
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	147..152
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	153..158
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	159..164
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	165..170
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	171..176
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	181..186
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	187..192
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	193..198
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	215..220
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	221..226
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	227..232
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	233..238
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	239..244
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	257..262
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	263..268
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	269..274
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	275..280
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	286..291
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	292..297
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	298..303
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	309..314
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	315..320
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif
FT	Peptide	322..328
FT		/label= GXYGX'Y' motif
FT		/note= "Claim 1" motif

```

FT Peptide /note= "Claim 1"
FT 329. .334 /label= GYGX'Y' motif
FT /note= "Claim 1"
FT Peptide 335. .340 /label= GYGX'Y' motif
FT /note= "Claim 1"
FT Peptide 354. .359 /label= GYGX'Y' motif
FT /note= "Claim 1"
FT Peptide 360. .365 /label= GYGX'Y' motif
FT /note= "Claim 1"
FT Peptide 366. .372 /label= GYGX'Y' motif
FT /note= "Claim 1"
FT Peptide 523. .528 /label= GYGX'Y' motif
FT /note= "Claim 1"
FT Peptide 542. .547 /label= GYGX'Y' motif
FT /note= "Claim 1"
FT Peptide 590. .595 /label= GYGX'Y' motif
FT /note= "Claim 1"
FT XX
FT PN US5643783-A.
FT XX
FT PD 01-JUL-1997.
FT XX
FT PF 01-DEC-1993; 93US-00159784.
FT XX
FT PR 01-DEC-1993; 93US-00159784.
FT XX
FT PA (HARD ) HARVARD COLLEGE.
FT XX
FT PI Olsen BR, Oh SP;
FT XX
FT DR WPI; 1997-350247/32.
FT N-PSDB; AAT84484.
FT XX
FT XX Nucleic acid encoding human alpha-1 collagen - for production of
FT recombinant alpha-1 collagen, for use in the treatment of cartilage
FT degeneration.
FT PS Claim 1; Col 23-30; 35pp; English.
FT XX
FT CC Novel human type alpha-1 (XVIII) collagen is characterised by 10 triple
FT helical domains containing the GYGX'Y' motif (where X, Y, X' and Y'
FT represent any amino acid), the helical domains being separated and
FT flanked by non-triple helical regions which may provide flexibility.
FT Alpha-1 collagen is expressed in multiple tissues, especially liver, lung
FT and kidney. A claimed plasmid comprising alpha-1 collagen nucleic acid
FT (see AAT84484) and an expression control sequence can be used to express
FT recombinant collagen in prokaryotic or eukaryotic (especially mammalian)
FT host cells. The alpha-1 collagen may be used to treat a patient suffering
FT from a disease associated with cartilage degradation, and for
FT supplementing collagen. It can also be used as a connective tissue filler
FT (e.g. for plastic surgery), can be interposed between a dermal equivalent
FT
FT Query Match 100.0%; Score 893; DB 2; Length 684;
FT Best Local Similarity 100.0%; Pred. No. 2.2e-100;
FT Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRAFLSSRLQDLYSIIVRRADRAAVPIV 60
DB 514 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRAFLSSRLQDLYSIIVRRADRAAVPIV 573
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPQKSVWHGSDPNGRRLTE 120
DB 574 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPQKSVWHGSDPNGRRLTE 633
QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170

```

```

DB 634 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 683
RESULT 30
AAAY25113
ID AAY25113 standard; protein; 684 AA.
XX AC AAY25113;
XX DT 25-AUG-1999 (first entry)
XX DE Human alpha1 (XVIII) collagen protein.
XX KW Alpha1(XVIII) collagen; mimetic; endostatin; atomic coordinate; library;
XX anti-angiogenic; heparin binding domain; receptor binding domain; mimic;
XX alpha-helix A domain; carbohydrate recognition domain; CRD domain;
XX treatment; angiogenesis; tumour; human.
XX OS Homo sapiens.
XX PN WO9931616-A1.
XX PD 24-JUN-1999.
XX PF 16-DEC-1998; 98WO-US026783.
XX PR 16-DEC-1997; 97US-0069727P.
XX PA (HARD ) HARVARD COLLEGE.
XX PI Olsen BR, Hohenester E, Timpl R, Sasaki T;
XX DR WPI; 1999-395243/33.
XX DR N-PSDB; AAX78379.
XX FT Identifying mimetics of mammalian endostatin.
XX PS Disclosure; Fig 5A-C; 75pp; English.
XX CC This invention describes a novel method for identifying mimetics of
XX mammalian endostatin. The method comprises identifying a compound having
XX atomic coordinates with non-trivial similarity to selected coordinates of
XX atoms of a mammalian endostatin involves (a) providing a library of
XX atomic coordinates of compounds in a library of candidate compounds, (b)
XX comparing the library of atomic coordinates to the selected coordinates
XX of a mammalian endostatin and (c) selecting from the library at least one
XX candidate compound on the basis of selection criteria which include
XX similarities between the atomic coordinates of the selected candidate
XX compound and the atomic coordinates of the mammalian endostatin. The
XX invention also describes the use of an anti-angiogenic fragment of
XX endostatin comprising a domain selected from a heparin binding domain, a
XX receptor binding domain, and exposed on alpha-helix A domain, and a
XX carbohydrate recognition domain (CRD) domain. The methods can be used for
XX designing and selecting endostatin mimics. The compounds identified can
XX be used for treating undesired angiogenesis, e.g. tumours. This sequence
XX represents human alpha1(XVIII) collagen which is used in the description
XX of the method
XX
XX Query Match 100.0%; Score 893; DB 2; Length 684;
XX Best Local Similarity 100.0%; Pred. No. 2.2e-100;
XX Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRAFLSSRLQDLYSIIVRRADRAAVPIV 60
DB 514 VALNSPLSGMGRGIRGADFCQFQARAVGLAGTFRAFLSSRLQDLYSIIVRRADRAAVPIV 573
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPQKSVWHGSDPNGRRLTE 120
DB 574 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPQKSVWHGSDPNGRRLTE 633

```



QY 121 SYCETWTEAPSGATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 DB 634 SYCETWTEAPSGATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 683

RESULT 31  
 AAO17357  
 ID AAO17357 standard; protein; 684 AA.  
 AC AAO17357;  
 XX 19-JUL-2002 (first entry)  
 XX Human collagen type XVIII alpha 1.  
 DE Human; endometriosis; DNA chip; fibronectin; p27; reticulocalbin;  
 XX aldehyde dehydrogenase 6; gravin; phospholipase C epsilon; elastin;  
 KW insulin-like growth factor binding protein-2; alpha-2 type IV collagen;  
 KW transmembrane receptor PTK7; collagen type XVIII alpha 1;  
 KW platelet derived growth factor receptor alpha; laminin M chain;  
 KW subtilisin like protein PACE4; nidogen.  
 XX Homo sapiens.  
 OS EP1191107-A2.  
 PN 27-MAR-2002.  
 PD 21-AUG-2001; 2001EP-00250300.  
 PF 25-SEP-2000; 2000DE-01048633.  
 PR (SCHD) SCHERING AG.  
 PA Hess-Stumpp H, Haendler B, Kraetzschmar J, Kreft B, Winterhager E;  
 PI Regidor P, Scotti S;  
 PI WPI; 2002-317413/36.  
 DR In vitro diagnosis and monitoring of endometriosis, comprises detecting  
 PT reduced expression of specific gene products, e.g. from the fibronectin  
 PT gene.  
 PS Claim 1; Page 12-13; 21pp; German.  
 XX The present invention relates to a method for the in vitro diagnosis of  
 CC endometriosis by determining the amount of gene product from at least one  
 CC specific gene in a patient sample and comparing this with the amount of  
 CC gene product in a control sample. A reduced level is indicative of  
 CC endometriosis. The gene products may be fibronectin, p27, reticulocalbin,  
 CC aldehyde dehydrogenase 6, gravin, phospholipase C epsilon, elastin,  
 CC insulin-like growth factor binding protein-2, alpha-2 type IV collagen,  
 CC transmembrane receptor PTK7, collagen type XVIII alpha 1, platelet  
 CC derived growth factor receptor alpha, laminin M chain, subtilisin like  
 CC protein PACE4 or nidogen. The method is useful for initial diagnosis of  
 CC endometriosis, and also for monitoring progress and treatment of the  
 CC disease. The present sequence is human collagen type XVIII alpha 1  
 XX  
 SQ Sequence 684 AA;

Query Match 100.0%; Score 893; DB 5; Length 684;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-100;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAPVIV 60  
 DB 514 VALNSPLSGMGRGIRGADFCQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAPVIV 573  
 QY 61 NLKDELLFSPSEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 574 NLKDELLFSPSEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 633  
 QY 121 SYCETWTEAPSGATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170

DB 634 SYCETWTEAPSGATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 683

RESULT 32  
 AAW92296  
 ID AAW92296 standard; peptide; 1301 AA.  
 AC AAW92296;  
 XX 28-APR-1999 (first entry)  
 XX Human alpha-1 (XVIII) collagen chain common sequence HU18 (common) 36.  
 DE Human; type XVIII collagen; liver disease; cirrhosis; detection;  
 XX hepatocellular carcinoma; diagnosis.  
 KW Homo sapiens.  
 OS WO9856399-A1.  
 PN 17-DEC-1998.  
 PD 12-JUN-1998; 98WO-US012327.  
 PF 12-JUN-1997; 97US-0049369P.  
 PR (FIBR-) FIBROGEN INC.  
 PA (FIFI-) ACAD FINLAND.  
 PA (INEM) INST NAT SANTE & RECH MEDICALE.  
 XX Pihlajaniemi T, Rehn M, Clement B;  
 PI WPI; 1999-070292/06.  
 DR Diagnosis and monitoring of liver disease by measuring collagen type  
 XX XVIII levels - with elevated levels indicative of disease, especially  
 PT cirrhosis or hepatocellular carcinoma.  
 PS Example 6; Fig 8; 56pp; English.  
 XX A method has been developed for the detecting liver disease. The method  
 CC comprises: (a) reacting a patient sample with antibodies (Ab) specific  
 CC for collagen type XVIII (Coll8); (b) measuring the amount of Ab-antigen  
 CC complex (C) formed as indicator of the amount of Coll8 present; (c)  
 CC similar analysis of a non-diseased control; and (d) comparing the amounts  
 CC of Coll8 in the two samples to detect presence or progression of disease.  
 CC Elevated levels of Coll8 are: (i) indicative of disease, specifically  
 CC cirrhosis; and (ii) predictive of the prognosis of disease, specifically  
 CC hepatocellular carcinoma (there is a relationship between Coll8 mRNA  
 CC levels and tumour size and necrosis, and survival times are significantly  
 CC higher in patients with higher Coll8 levels). The method provides non-  
 CC invasive, early and accurate diagnosis of liver disease. The present  
 CC sequence represents the sequence common to human alpha-1 (XVIII) collagen  
 CC chain from the present invention  
 XX  
 SQ Sequence 1301 AA;

Query Match 100.0%; Score 893; DB 2; Length 1301;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-100;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAPVIV 60  
 DB 1131 VALNSPLSGMGRGIRGADFCQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAPVIV 1190  
 QY 61 NLKDELLFSPSEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 120  
 DB 1191 NLKDELLFSPSEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 1250  
 QY 121 SYCETWTEAPSGATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 DB 1251 SYCETWTEAPSGATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 1300

XX AC ABP96308;  
XX DT 20-MAY-2003 (first entry)  
XX DE Human endostatin protein.  
XX KW Humanised baculovirus; cytostatic; gene therapy; baculovirus; cancer;  
XX KW prostate cancer; endostatin; chromosome 21.  
XX OS Homo sapiens.  
XX PN WO2003016540-A2.  
XX PD 27-FEB-2003.  
XX PF 15-AUG-2002; 2002WO-GB003791.  
XX PR 15-AUG-2001; 2001GB-00019852.  
XX PA (UYO-) UNIV YORK.  
XX PI Maitland N;  
XX DR WPI; 2003-268336/26.  
XX PT New baculovirus having a modified genome encoding a therapeutic agent,  
XX PT useful in the manufacture of a medicament for the treatment of cancer,  
XX PT particularly prostate cancer.  
XX PS Claim 24; Page; 34pp; English.  
XX CC The present invention describes a humanised baculovirus (I) which  
XX CC comprises a modified baculovirus genome having a nucleic acid molecule  
XX CC encoding a therapeutic agent and a polypeptide which functions to target  
XX CC the baculovirus to at least one cell type. Also described is a  
XX CC pharmaceutical composition comprising (I). (I) has cytostatic activity,  
XX CC and can be used in gene therapy. The baculovirus is useful in the  
XX CC manufacture of a medicament for the treatment of cancer, particularly  
XX CC prostate cancer. The present sequence represents the human endostatin  
XX CC protein, which is specified in the exemplification of the present  
XX CC invention. N.B. The present sequence is not given in the specification  
XX CC but is referred to in Claim 24 as Genbank accession number NM\_130445  
XX SQ Sequence 1336 AA;  
Query Match 100.0%; Score 893; DB 6; Length 1336;  
Best Local Similarity 100.0%; Pred. No. 6e-100;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60  
Db 1166 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 1225  
QY 61 NLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVLHPTWPKSVWHGSDPNRRLTE 120  
Db 1226 NLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVLHPTWPKSVWHGSDPNRRLTE 1285  
QY 121 SYCETWRTEAPSATQASLLGRLGQSAASHAYIVLCIENSFMTAS 170  
Db 1286 SYCETWRTEAPSATQASLLGRLGQSAASHAYIVLCIENSFMTAS 1335  
RESULT 35  
ABP96308  
ID ABB83471 standard; protein; 1516 AA.  
XX AC ABB83471;  
XX DT 30-SEP-2002 (first entry)  
XX DE Human collagen XVIII.  
XX

RESULT 33  
RAY08694  
ID AAY08694 standard; protein; 1336 AA.  
XX AC AAY08694;  
XX DT 10-AUG-1999 (first entry)  
XX DE Human collagen 18 protein.  
XX KW Plasminogen; human; angiotensin; endostatin; gene therapy; vector;  
XX KW anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;  
XX KW tumour growth; solid tumour; diabetic retinopathy; retina; collagen 18.  
XX OS Homo sapiens.  
XX PN WO9926480-A1.  
XX PD 03-JUN-1999.  
XX PF 20-NOV-1998; 98WO-US024950.  
XX PR 20-NOV-1997; 97US-00975424.  
XX PA (GENE-) GENETIX PHARM INC.  
XX PA (MASI) MASSACHUSETTS INST TECHNOLOGY.  
XX PI Le Boulch P, Pawliuk RJ, Bachelot T;  
XX DR WPI; 1999-357696/30.  
XX DR N-PSDB; AAX77720.  
XX PT Anti-angiogenic gene therapy vectors.  
XX PS Disclosure; Page 77-80; 83pp; English.  
XX CC This invention describes a novel viral gene therapy vector comprising a  
XX CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen from  
XX CC human or murine angiotensin, human or murine endostatin and angiotensin-  
XX CC inhibiting fusions and fragments, where the viral vector is sufficiently  
XX CC attenuated for use in human gene therapy. The products of the invention  
XX CC have anti-angiogenic, cytostatic, anti-diabetic and ophthalmological  
XX CC activity. The vector is used in gene therapy for inhibiting tumour growth  
XX CC in humans harbouring a solid tumour. The vector expresses an anti-  
XX CC angiogenic polypeptide. An additional use comprises treatment of diabetic  
XX CC retinopathy, where the anti-angiogenic polypeptide inhibits angiogenesis  
XX CC in the vicinity of the retina. The vector is administered to cells ex  
XX CC vivo and then administered to the patient  
XX SQ Sequence 1336 AA;  
Query Match 100.0%; Score 893; DB 2; Length 1336;  
Best Local Similarity 100.0%; Pred. No. 6e-100;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60  
Db 1166 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 1225  
QY 61 NLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVLHPTWPKSVWHGSDPNRRLTE 120  
Db 1226 NLKDELLFPSWEALFSGSEGLKPGARIFSGDKDVLHPTWPKSVWHGSDPNRRLTE 1285  
QY 121 SYCETWRTEAPSATQASLLGRLGQSAASHAYIVLCIENSFMTAS 170  
Db 1286 SYCETWRTEAPSATQASLLGRLGQSAASHAYIVLCIENSFMTAS 1335  
RESULT 34  
ABP96308  
ID ABB96308 standard; protein; 1336 AA.

KW Human; antirheumatic; antiarthritic; gene therapy; anti-angiogenic;  
 KW rheumatoid arthritis; collagen; endostatin.  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 1334.1516  
 PT Protein /note= "This region is specifically claimed in Claim 4"  
 FT  
 XX  
 XX WO200253191-A1.  
 PN  
 XX  
 XX 11-JUL-2002.  
 PD  
 XX  
 XX 03-JAN-2002; 2002WO-KR000001.  
 PF  
 XX  
 XX 05-JAN-2001; 2001KR-00000691.  
 PR  
 XX  
 XX (VIRO-) VIROMED LTD.  
 PA  
 XX  
 XX Kim J, Ho S, Park E, Kim S;  
 PI  
 XX  
 XX WPI; 2002-583596/62.  
 DR  
 XX  
 XX N-PSDB; ABN85301.  
 DR  
 XX  
 XX Novel composition for gene therapy against rheumatoid arthritis,  
 PT comprising a DNA encoding anti-angiogenic protein or its parts.  
 PT  
 XX  
 XX Claim 4; Page 70-78; 84pp; English.  
 PS  
 XX  
 XX The present invention relates to a composition for gene therapy,  
 CC comprising a DNA encoding an anti-angiogenic protein, which shows  
 CC therapeutic effects on rheumatoid arthritis. The composition is useful  
 CC for treating rheumatoid arthritis and the gene therapy is effective,  
 CC lasting for 14 days. The present sequence is the protein sequence for  
 CC human collagen XVIII. Endostatin, which consists of the C-terminal 183  
 CC residues of collagen XVIII, was used as an anti-angiogenic protein  
 XX  
 XX Sequence 1516 AA;  
 SQ  
 Query Match 100.0%; Score 893; DB 5; Length 1516;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-100;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCQFQQARAVGLAGTFFAFLSSRLQDLYSIVRRADRAAVPIV 60  
 DB 1346 VALNSPLSGMGRGIRGADFCQFQQARAVGLAGTFFAFLSSRLQDLYSIVRRADRAAVPIV 1405  
 QY 61 NLKDELLFPSEALFSGSEGPLKPGARIFSPDGKDLRHPTWPKSVWHGSDPNGRRLTE 120  
 DB 1406 NLKDELLFPSEALFSGSEGPLKPGARIFSPDGKDLRHPTWPKSVWHGSDPNGRRLTE 1465  
 QY 121 SYCETWRTAPSATQASLLGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 DB 1466 SYCETWRTAPSATQASLLGRLGQSAASCHHAYIVLCIENSFMTAS 1515  
 RESULT 36  
 ABP68617  
 ID ABP68617 standard; protein; 1516 AA.  
 XX  
 XX AC ABP68617;  
 XX  
 XX DT 14-JAN-2003 (first entry)  
 XX  
 XX Human pancreatic cancer expressed protein SEQ ID NO 166.  
 DE  
 XX  
 XX Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;  
 KW  
 KW cytotatic; tumour.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200260317-A2.  
 PN  
 XX

08-AUG-2002.  
 30-JAN-2002; 2002WO-US002781.  
 30-JAN-2001; 2001US-0265305P.  
 31-JAN-2001; 2001US-0265882P.  
 09-FEB-2001; 2001US-0267568P.  
 21-MAR-2001; 2001US-0278651P.  
 28-APR-2001; 2001US-0287112P.  
 16-MAY-2001; 2001US-0291631P.  
 12-JUL-2001; 2001US-0305484P.  
 20-AUG-2001; 2001US-0313999P.  
 27-NOV-2001; 2001US-0333626P.  
 (CORI-) CORIXA CORP.  
 Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;  
 WPI; 2002-627435/67.  
 N-PSDB; ABV94763.  
 New isolated polynucleotide and pancreatic tumor polypeptides, useful for  
 diagnosing, preventing and/or treating cancer, particularly pancreatic  
 cancer.  
 Claim 2; SEQ ID NO 166; 300pp + Sequence Listing; English.  
 The invention relates to an isolated polynucleotide (I) comprising: (a)  
 any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)  
 complements of (a); (c) sequences consisting of at least 20 contiguous  
 residues of (a); (d) sequences that hybridize to (a), under moderately  
 stringent conditions; (e) sequences having at least 75% or 90% identity  
 to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-  
 ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer  
 in a patient and compositions comprising polypeptides, polynucleotides,  
 antibodies, fusion proteins, T cell populations and antigen presenting  
 cells expressing the polypeptide are useful in treating pancreatic cancer  
 and stimulating an immune response. The polynucleotides can be used as  
 probes or primers for nucleic acid hybridization, in the design and  
 preparation of ribozyme molecules for inhibiting expression of the tumour  
 polypeptides and proteins in the tumour cells, in vaccines and for gene  
 therapy. Note: The sequence data for this patent did not form part of the  
 printed specification, but was obtained in electronic format directly  
 from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 Sequence 1516 AA;  
 Query Match 100.0%; Score 893; DB 5; Length 1516;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-100; Mismatches 0; Gaps 0;  
 Matches 170; Conservative 0; Indels 0; Gaps 0;  
 QY 1 VALNSPLSGMGRGIRGADFCQFQQARAVGLAGTFFAFLSSRLQDLYSIVRRADRAAVPIV 60  
 DB 1346 VALNSPLSGMGRGIRGADFCQFQQARAVGLAGTFFAFLSSRLQDLYSIVRRADRAAVPIV 1405  
 QY 61 NLKDELLFPSEALFSGSEGPLKPGARIFSPDGKDLRHPTWPKSVWHGSDPNGRRLTE 120  
 DB 1406 NLKDELLFPSEALFSGSEGPLKPGARIFSPDGKDLRHPTWPKSVWHGSDPNGRRLTE 1465  
 QY 121 SYCETWRTAPSATQASLLGRLGQSAASCHHAYIVLCIENSFMTAS 170  
 DB 1466 SYCETWRTAPSATQASLLGRLGQSAASCHHAYIVLCIENSFMTAS 1515  
 Search completed: March 13, 2004, 08:29:11  
 JOB time : 59 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 08:17:57 ; Search time 34 Seconds  
(without alignment)

1055.766 Million cell updates/sec

Title: US-09-171-607A-1

Perfect score: 893

Sequence: 1 VALNSPLSGMGRGADFCQ.....ASCHHAVILCIENSFMFAS 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/ECTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	893	100.0	178	14	US-10-131-241-60
2	893	100.0	178	14	Sequence 60, Appl
3	893	100.0	179	14	Sequence 5, Appl
4	893	100.0	179	14	Sequence 57, Appl
5	893	100.0	182	9	US-10-131-241-57
6	893	100.0	182	14	Sequence 14, Appl
7	893	100.0	182	14	Sequence 54, Appl
8	893	100.0	182	14	Sequence 3, Appl
9	893	100.0	182	14	Sequence 14, Appl
10	893	100.0	183	13	US-10-373-561-14
11	893	100.0	183	13	US-09-873-676-2
12	893	100.0	183	13	US-10-080-797-1
13	893	100.0	183	14	US-10-131-241-52
14	893	100.0	183	14	Sequence 4, Appl
15	893	100.0	183	14	Sequence 3010, Ap
16	893	100.0	682	15	US-10-264-049-3010
17	893	100.0	684	10	US-09-961-403-5
18	893	100.0	1516	14	US-10-060-036-166
19	893	100.0	1516	15	US-10-431-642-3

16	889	99.6	181	14	US-10-131-241-55	Sequence 55, Appl
17	885	99.1	180	14	US-10-131-241-56	Sequence 56, Appl
18	858	96.1	180	14	US-10-131-241-47	Sequence 47, Appl
19	783	87.7	184	10	US-09-938-391-4	Sequence 4, Appl
20	783	87.7	184	14	US-10-131-241-49	Sequence 49, Appl
21	783	87.7	184	14	US-10-292-418-35	Sequence 35, Appl
22	783	87.7	230	10	US-09-938-391-2	Sequence 2, Appl
23	778	87.1	191	9	US-09-988-831-13	Sequence 13, Appl
24	778	87.1	191	14	US-10-373-561-13	Sequence 13, Appl
25	778	87.1	207	13	US-10-080-797-3	Sequence 3, Appl
26	775	86.8	184	14	US-10-131-241-46	Sequence 46, Appl
27	775	86.8	184	14	US-10-292-418-18	Sequence 18, Appl
28	775	86.8	207	14	US-10-422-934-71	Sequence 71, Appl
29	734	82.2	185	13	US-10-036-869-36	Sequence 36, Appl
30	346	38.7	63	9	US-09-822-540A-1	Sequence 1, Appl
31	182	20.4	31	9	US-09-822-540A-2	Sequence 2, Appl
32	109	12.2	24	15	US-10-016-569A-29	Sequence 29, Appl
33	109	12.2	24	15	US-10-308-644-29	Sequence 29, Appl
34	94	10.5	16	9	US-09-766-412-32	Sequence 32, Appl
35	85	9.5	1244	9	US-09-815-915-8	Sequence 8, Appl
36	85	9.5	1244	14	US-10-393-316-8	Sequence 8, Appl
37	80	9.0	409	14	US-10-156-761-13624	Sequence 13624, A
38	78.5	8.8	285	15	US-10-369-493-12038	Sequence 12038, A
39	77.5	8.7	418	9	US-09-927-602-5	Sequence 5, Appl
40	77	8.6	451	15	US-10-369-493-9525	Sequence 9525, Ap
41	77	8.6	1267	10	US-09-769-734-38	Sequence 38, Appl
42	76.5	8.6	421	14	US-10-156-761-13612	Sequence 13612, A
43	76	8.5	239	15	US-10-369-493-8031	Sequence 8031, Ap
44	76	8.5	636	15	US-10-351-157-9	Sequence 9, Appl
45	75.5	8.5	3192	14	US-10-132-134-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

- US-10-131-241-60
- Sequence 60, Application US/10131241
- Publication No. US20030012792A1
- GENERAL INFORMATION:
- APPLICANT: Holaday, John W.
- TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolife
- TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
- FILE REFERENCE: 05213-0344 43170-271565
- CURRENT APPLICATION NUMBER: US/10/131,241
- CURRENT FILING DATE: 2002-07-22
- PRIOR APPLICATION NUMBER: US 09/413,049
- PRIOR FILING DATE: 1993-10-06
- PRIOR APPLICATION NUMBER: US 09/316,802
- PRIOR FILING DATE: 1999-05-21
- PRIOR APPLICATION NUMBER: US 60/086,586
- PRIOR FILING DATE: 1998-05-22
- NUMBER OF SEQ ID NOS: 65
- SOFTWARE: PatentIn version 3.1
- SEQ ID NO 60
- LENGTH: 178
- TYPE: PRT
- ORGANISM: Homo sapiens
- US-10-131-241-60

Query Match 100.0%; Score 893; DB 14; Length 178;  
Best Local Similarity 100.0%; Pred. No. 3.4e-92;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VALNSPLSGMGRGADFCQFQQAQAVGLAGTFRFLSSRLQDLYSIVRRADRAAPVIV 60
Db	9	VALNSPLSGMGRGADFCQFQQAQAVGLAGTFRFLSSRLQDLYSIVRRADRAAPVIV 68
Qy	61	NLKDELLFFSWAELFSGSEGLPKPGARIIFSFGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
Db	69	NLKDELLFFSWAELFSGSEGLPKPGARIIFSFGKDVLRHPTWPKSVWHGSDPNGRRLTE 128

```
QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in version 3.1
Db 129 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 178
; SEQ ID NO 57
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-131-241-57

Query Match 100.0%; Score 893; DB 14; Length 179;
Best Local Similarity 100.0%; Pred. No. 3.4e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
US-10-042-347-5
; Sequence 5, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fu
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-042-347-5

Query Match 100.0%; Score 893; DB 14; Length 178;
Best Local Similarity 100.0%; Pred. No. 3.4e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGGMRGIRGADFCQFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60
;
Db 9 VALNSPLSGGMRGIRGADFCQFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 68
;
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
;
Db 69 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 128
;
QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170
;
Db 129 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 178
;
US-10-131-241-57

Query Match 100.0%; Score 893; DB 9; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.5e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
US-10-131-241-57
; Sequence 57, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
```

```
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 57
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-131-241-57

Query Match 100.0%; Score 893; DB 14; Length 179;
Best Local Similarity 100.0%; Pred. No. 3.4e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGGMRGIRGADFCQFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60
;
Db 9 VALNSPLSGGMRGIRGADFCQFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 68
;
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
;
Db 69 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 128
;
QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170
;
Db 129 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 178
;
US-09-998-831-14

RESULT 4
US-09-998-831-14
; Sequence 14, Application US/09998831
; Patent No. US20020119153A1
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY
; TITLE OF INVENTION: INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/998,831
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/561,108
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-998-831-14

Query Match 100.0%; Score 893; DB 9; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.5e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGGMRGIRGADFCQFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60
;
Db 13 VALNSPLSGGMRGIRGADFCQFQQAARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 72
;
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
;
Db 73 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 132
;
QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170
;
Db 133 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 182
;
US-10-131-241-54

RESULT 5
US-10-131-241-54
; Sequence 54, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-241-54

Query Match 100.0%; Score 893; DB 14; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.5e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGADGFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVPIV 60
DB 13 VALNSPLSGMGRGADGFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVPIV 72

QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
DB 73 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 6
US-10-042-347-3
; Sequence 3, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide F
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-347-3

Query Match 100.0%; Score 893; DB 14; Length 182;
```

```
Best Local Similarity 100.0%; Pred. No. 3.5e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGADGFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVPIV 60
DB 13 VALNSPLSGMGRGADGFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVPIV 72

QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
DB 73 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 7
US-10-373-561-14
; Sequence 14, Application US/10373561
; Publication No. US20030175276A1
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002582
; CURRENT APPLICATION NUMBER: US/10/373,561
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US/09/561,499
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-10-373-561-14

Query Match 100.0%; Score 893; DB 14; Length 182;
Best Local Similarity 100.0%; Pred. No. 3.5e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGADGFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVPIV 60
DB 13 VALNSPLSGMGRGADGFCQFQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVPIV 72

QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120
DB 73 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 132

QY 121 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
DB 133 SYCETWRTAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 8
US-09-873-676-2
; Sequence 2, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: Macdonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiotensin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
```

us-09-171-607a-1.rapb

Sat Mar 13 08:24:05 2004

```

; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-131-241-52

Query Match 100.0%; Score 893; DB 14; Length 183;
Best Local Similarity 100.0%; Pred. No. 3.5e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VALNSPLSGMGRGIRGADFCFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60
Db 13 VALNSPLSGMGRGIRGADFCFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 72
Qy 61 NLKDELLFPPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 120
Db 73 NLKDELLFPPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 132
Qy 121 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
Db 133 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 9
US-10-080-797-1
; Sequence 1, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campochiaro, Peter A.
; APPLICANT: Dixon, Katharine H.
; APPLICANT: Brazzell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; TITLE OF INVENTION: NEOVASCULARIZATION
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Human
; ORGANISM: Human
US-10-080-797-1

Query Match 100.0%; Score 893; DB 13; Length 183;
Best Local Similarity 100.0%; Pred. No. 3.5e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VALNSPLSGMGRGIRGADFCFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60
Db 13 VALNSPLSGMGRGIRGADFCFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 72
Qy 61 NLKDELLFPPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 120
Db 73 NLKDELLFPPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 132
Qy 121 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
Db 133 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 10
US-10-131-241-52
; Sequence 52, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers

```

```

; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-131-241-52

Query Match 100.0%; Score 893; DB 14; Length 183;
Best Local Similarity 100.0%; Pred. No. 3.5e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VALNSPLSGMGRGIRGADFCFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60
Db 13 VALNSPLSGMGRGIRGADFCFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 72
Qy 61 NLKDELLFPPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 120
Db 73 NLKDELLFPPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 132
Qy 121 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170
Db 133 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 11
US-10-292-418-4
; Sequence 4, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-292-418-4

Query Match 100.0%; Score 893; DB 14; Length 183;
Best Local Similarity 100.0%; Pred. No. 3.5e-92;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VALNSPLSGMGRGIRGADFCFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 60
Db 13 VALNSPLSGMGRGIRGADFCFQOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAIVP 72
Qy 61 NLKDELLFPPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 120
Db 73 NLKDELLFPPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNRRLTE 132
Qy 121 SYCETWTEAPSATGQASSLLGRLGQSAASCHHAYIVLCIENSFMTAS 170

```

Db 133 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 182

## RESULT 12

US-10-264-049-3010  
; Sequence 3010, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; PRIOR FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: US 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 3010  
; LENGTH: 682  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (20)\_FEATURE  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (39)  
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
US-10-264-049-3010

Query Match 100.0%; Score 893; DB 15; Length 682;  
Best Local Similarity 100.0%; Pred. No. 2e-91;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 60  
Db 512 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 571  
  
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
Db 572 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 631  
  
QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
Db 632 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 681

## RESULT 13

US-09-961-403-5  
; Sequence 5, Application US/09961403  
; Publication No. US20030077589A1  
; GENERAL INFORMATION:  
; APPLICANT: HE-STUMPP, HOLGER  
; APPLICANT: HAENDLER, BERNARD  
; APPLICANT: KRAETZSCHWAR, JOERN  
; APPLICANT: KREFT, BERTHOLT  
; APPLICANT: REGIDOR, PEDRO  
; APPLICANT: SCOTTI, SIMONE  
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS  
; FILE REFERENCE: SCH-1789  
; CURRENT APPLICATION NUMBER: US/09/961,403  
; CURRENT FILING DATE: 2001-09-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 684  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-961-403-5

Query Match 100.0%; Score 893; DB 10; Length 684;  
Best Local Similarity 100.0%; Pred. No. 2e-91;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 60  
Db 514 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 573  
  
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
Db 574 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 633  
  
QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
Db 634 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 683

## RESULT 14

US-10-060-036-166  
; Sequence 166, Application US/10060036  
; Publication No. US20030073144A1  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Persing, David H.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Jiang, Yuqiu  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
; FILE REFERENCE: 210121.566  
; CURRENT APPLICATION NUMBER: US/10/060,036  
; CURRENT FILING DATE: 2002-01-30  
; NUMBER OF SEQ ID NOS: 4560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 166  
; LENGTH: 1516  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-060-036-166

Query Match 100.0%; Score 893; DB 14; Length 1516;  
Best Local Similarity 100.0%; Pred. No. 5.6e-91;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 60  
Db 1346 VALNSPLSGMGRGIRGADFCFQOARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP 1405  
  
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
Db 1406 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 1465  
  
QY 121 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
Db 1466 SYCETWTEAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 1515

## RESULT 15

US-10-431-642-3  
; Sequence 3, Application US/10431642  
; Publication No. US20040009920A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING  
; TITLE OF INVENTION: TUMOR GROWTH AND ANGIOGENESIS  
; FILE REFERENCE: BURNHAM.008CIP  
; CURRENT APPLICATION NUMBER: US/10/431,642  
; CURRENT FILING DATE: 2003-05-05  
; PRIOR APPLICATION NUMBER: 10/005,171  
; PRIOR FILING DATE: 2001-12-03  
; PRIOR APPLICATION NUMBER: 60/331,357



us-09-171-607a-1.rapb

Sat Mar 13 08:24:05 2004

```

; PRIOR FILING DATE: 2000-12-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1516
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-431-642-3

Query Match      100.0%; Score 893; DB 15; Length 1516;
Best Local Similarity 100.0%; Pred. No. 5.6e-91;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VALNSPLSGMGRGIRGADFCQCFQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 60
Db      1346 VALNSPLSGMGRGIRGADFCQCFQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 1405

QY      61 NLKDELLFESWEALFSGSGPLKPGARIFFDQKDVLRHPTWPKSVWHGSDPNGRRLTE 120
Db      1406 NLKDELLFESWEALFSGSGPLKPGARIFFDQKDVLRHPTWPKSVWHGSDPNGRRLTE 1465

QY      121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWTAS 170
Db      1466 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFWTAS 1515

```

Search completed: March 13, 2004, 08:23:27  
Job time : 35 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 08:15:22 ; Search time 22 Seconds  
(without alignments)  
398.928 Million cell updates/sec

Title: US-09-171-607A-1  
Perfect score: 893  
Sequence: 1 VALNSPLSGMGRGADPQ.....ASCHHAYIVLCIENSFWTAS 170

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues  
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pap.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pap.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pap.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pap.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pap.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	893	100.0	178	4	US-09-315-689-5
2	893	100.0	182	4	US-09-561-500-14
3	893	100.0	182	4	US-09-561-108-14
4	893	100.0	182	4	US-09-315-689-3
5	893	100.0	182	4	US-09-561-526-14
6	893	100.0	182	4	US-09-561-499-14
7	893	100.0	182	4	US-09-998-831-14
8	893	100.0	183	3	US-09-206-059-2
9	778	87.1	191	4	US-09-561-500-13
10	778	87.1	191	4	US-09-561-108-13
11	778	87.1	191	4	US-09-561-526-13
12	778	87.1	191	4	US-09-561-499-13
13	778	87.1	191	4	US-09-998-831-13
14	775	86.8	195	1	US-08-159-784-2
15	734	82.2	185	3	US-08-985-526-36
16	491	55.0	191	1	US-08-159-784-3
17	242.5	27.2	123	4	US-09-231-077D-11
18	214.5	24.0	124	4	US-09-231-077D-10
19	160	17.9	35	3	US-09-046-985-2
20	160	17.9	35	3	US-09-474-743-2
21	101	11.3	22	3	US-09-046-985-7
22	101	11.3	22	3	US-09-474-743-7
23	94	10.5	16	3	US-09-385-442-32
24	76	8.5	403	4	US-08-252-991A-22238
25	76	8.5	578	1	US-08-653-740-3
26	76	8.5	578	2	US-09-073-594-3
27	76	8.5	578	3	US-09-275-925-3

28	76	8.5	636	1	US-08-653-740-5	Sequence 5, Appli
29	76	8.5	636	2	US-09-073-594-5	Sequence 5, Appli
30	76	8.5	636	3	US-09-275-925-5	Sequence 5, Appli
31	74	8.3	256	1	US-07-306-349A-8	Sequence 8, Appli
32	74	8.3	256	1	US-08-167-035-4	Sequence 4, Appli
33	74	8.3	256	1	US-08-208-887A-4	Sequence 4, Appli
34	74	8.3	256	2	US-08-539-005-4	Sequence 4, Appli
35	74	8.3	256	4	US-09-280-598-8	Sequence 8, Appli
36	73	8.2	311	3	US-08-987-743-2	Sequence 1, Appli
37	73	8.2	435	3	US-08-733-360A-1	Sequence 1, Appli
38	73	8.2	435	3	US-08-733-360A-3	Sequence 6, Appli
39	73	8.2	435	3	US-08-987-743-6	Sequence 15, Appli
40	73	8.2	435	3	US-08-987-743-15	Sequence 1, Appli
41	73	8.2	435	3	US-08-916-935-1	Sequence 3, Appli
42	73	8.2	435	3	US-08-916-935-3	Sequence 19779, A
43	72.5	8.1	184	4	US-09-252-991A-19779	Sequence 28, Appli
44	72	8.1	304	4	US-08-630-315A-28	Sequence 12, Appli
45	72	8.1	1006	4	US-09-023-905A-12	

ALIGNMENTS

RESULT 1  
US-09-315-689-5  
; Sequence 5, Application US/09315689  
; Patent No. 6346510  
; GENERAL INFORMATION:  
; APPLICANT: Folkman, Judah  
; APPLICANT: O'Reilly, Michael  
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions  
; FILE REFERENCE: 05213-0229  
; CURRENT APPLICATION NUMBER: US/09/315,689  
; CURRENT FILING DATE: 1999-05-20  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 178  
; TYPE: PXT  
; ORGANISM: Homo sapiens  
US-09-315-689-5

Query Match	100.0%;	Score	893;	DB	4;	Length	178;
Best Local Similarity	100.0%;	Pred. No.	4.5e-101;				
Matches	170;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	VALNSPLSGMGRGADPQCFQOARAVGLACTFRFLSSRLQDLYSIVRRADRAAPIV	60				
DB	9	VALNSPLSGMGRGADPQCFQOARAVGLACTFRFLSSRLQDLYSIVRRADRAAPIV	68				
QY	61	NLKDELLPPSWALFSGSGPLKPGARIFSDGKDLRHTWPQKSVHSGDPNGRRLTE	120				
DB	69	NLKDELLPPSWALFSGSGPLKPGARIFSDGKDLRHTWPQKSVHSGDPNGRRLTE	128				
QY	121	SYCETWRTAPSATGQASSLLGRLIGQSAASCHHAYIVLCIENSFWTAS	170				
DB	129	SYCETWRTAPSATGQASSLLGRLIGQSAASCHHAYIVLCIENSFWTAS	178				

RESULT 2  
US-09-561-500-14  
; Sequence 14, Application US/09561500  
; Patent No. 6342219  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brecken  
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002500  
; CURRENT APPLICATION NUMBER: US/09/561,500  
; CURRENT FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44

```

; APPLICANT: Folkman, Judah
;
; APPLICANT: O'Reilly, Michael
;
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
;
; FILE REFERENCE: 05213-0229
;
; CURRENT APPLICATION NUMBER: US/09/315,689
;
; CURRENT FILING DATE: 1999-05-20
;
; NUMBER OF SEQ ID NOS: 6
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 3
;
; LENGTH: 182
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
;
; RS-09-315-689-3

```

	Query Match	100.0%;	Score 893;	DB 4;	Length 182;
	Best Local Similarity	100.0%;	Fred. No. 4.6e-101;		
	Matches 170;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	VALNSPLSGGMGIRGADFCQFOQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP	60		
Db	13	VALNSPLSGGMGIRGADFCQFOQARAVGLAGTFFRAFLSSRLQDLYSIVRRADRAAIVP	72		
QY	61	NLKDELLFFSWEALFSGSGEPLKPGARIISFDGKDVLRHETWTPQKSVVHGSDPNGRLTE	120		
Db	73	NLKDELLFFSWEALFSGSGEPLKPGARIISFDGKDVLRHETWTPQKSVVHGSDPNGRLTE	132		
QY	121	SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAVIVLCIENSFMATAS	170		
Db	133	SYCETWRTAPSATGQASSLLGGRLGQSAASCHHAVIVLCIENSFMATAS	182		

```

/ Sequence # 442221
/ Patent No. 634221
/ GENERAL INFORMATION:
/ APPLICANT: Philip E. Thorpe
/ APPLICANT: Rolf A. Brekken
/ TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
/ FILE REFERENCE: 4001.002584
/ CURRENT APPLICATION NUMBER: US/09/561,108
/ CURRENT FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/131,432
/ PRIOR FILING DATE: 1999-04-28
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 14
/ LENGTH: 192
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
/ OTHER INFORMATION: PEPTIDE
/ 1S-09-561-108-14

```

Query Match	100.0%	Score	893	DB	4	Length	182
Best Local Similarity	100.0%	Prod. No.	4.ee-101	Indels	0	Gaps	0
Matches	170	Conservative	0	Mismatches	0		
QY	1	VALNSPLSGGMRGIGADFQCFQQARAVGLAGTFFAFLSSRLQDIYSIVRRADRRAVPV	60				
DB	13	VALNSPLSGGMRGIGADFQCFQQARAVGLAGTFFAFLSSRLQDIYSIVRRADRRAVPV	72				
QY	61	NLKDELLPSPNEALFSGSEGPLKPGCARIFSPGKDVLRHPTWPKQSVMHGSDPNGRRLTE	120				
DB	73	NLKDELLPSPNEALFSGSEGPLKPGCARIFSPGKDVLRHPTWPKQSVMHGSDPNGRRLTE	132				
QY	121	SYCETWRTEAPSATQGSALLGRLIGQSAASCHHAYIVLCIENSFWTAS	170				
DB	133	SYCETWRTEAPSATQGSALLGRLIGQSAASCHHAYIVLCIENSFWTAS	182				
RESULT 4							

US-09-315-689-3  
; Sequence 3, Application US/09315689  
; Patent No. 6346510

## RESULT 6

US-09-561-499-14  
; Sequence 14, Application US/09561499  
; Patent No. 6524583  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002582  
; CURRENT APPLICATION NUMBER: US/09/561,499  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
; OTHER INFORMATION: PEPTIDE  
US-09-561-499-14

Query Match 100.0%; Score 893; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 4.6e-101; Indels 0; Gaps 0;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 60  
DB 13 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 72  
QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 132  
QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
DB 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 8  
US-09-206-059-2  
; Sequence 2, Application US/09206059  
; Patent No. 6201104  
; GENERAL INFORMATION:  
; APPLICANT: MacDonald, Nicholas  
; APPLICANT: Sim, Kim Lee  
; TITLE OF INVENTION: Angiogenesis-Inhibiting Protein Binding Peptides and  
; TITLE OF INVENTION: Proteins and Methods of Use  
; FILE REFERENCE: 05213-0370  
; CURRENT APPLICATION NUMBER: US/09/206,059  
; CURRENT FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 183  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-206-059-2

Query Match 100.0%; Score 893; DB 3; Length 183;  
Best Local Similarity 100.0%; Pred. No. 4.7e-101; Indels 0; Gaps 0;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 60  
DB 13 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 72  
QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 132  
QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
DB 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 9  
US-09-561-500-13  
; Sequence 13, Application US/09561500  
; Patent No. 6342219  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002500  
; CURRENT APPLICATION NUMBER: US/09/561,500  
; CURRENT FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-561-500-13

Query Match 87.1%; Score 778; DB 4; Length 191;  
Best Local Similarity 85.8%; Pred. No. 5.6e-87; Indels 0; Gaps 0;  
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

US-09-561-499-14  
; Sequence 14, Application US/09561499  
; Patent No. 6524583  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002582  
; CURRENT APPLICATION NUMBER: US/09/561,499  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
; OTHER INFORMATION: PEPTIDE  
US-09-561-499-14

Query Match 100.0%; Score 893; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 4.6e-101; Indels 0; Gaps 0;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 60  
DB 13 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 72  
QY 61 NLKDELLFPSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 120  
DB 73 NLKDELLFPSWEALFSGSEGPLKPGARIFSDGKDVLRHPTWPKSVWHGSDPNGRRLTE 132  
QY 121 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 170  
DB 133 SYCETWRTAPSATGQASSLLGGRLLGQSAASCHHAYIVLCIENSFMTAS 182

RESULT 7  
US-09-998-831-14  
; Sequence 14, Application US/09998831  
; Patent No. 6676941  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY  
; TITLE OF INVENTION: INHIBITING VEGF  
; FILE REFERENCE: 4001.002584  
; CURRENT APPLICATION NUMBER: US/09/998,831  
; CURRENT FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: 09/561,108  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
; OTHER INFORMATION: PEPTIDE  
US-09-998-831-14

Query Match 100.0%; Score 893; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 4.6e-101; Indels 0; Gaps 0;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 60  
DB 13 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRFLSSRLQDLYSIVRRADRAAVPIV 72

Sat Mar 13 08:24:05 2004

OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-561-526-13

Query Match 87.1%; Score 778; DB 4; Length 191;  
Best Local Similarity 85.8%; Pred. No. 5.6e-87;  
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;  
QY 1 VALNSPLSGMGRGIRGADFCQFQOARAVGLAGTFRALSSRLQDLYSIVRRADRAAIVP 60  
DB 20 VALNTPLSGGMRGIRGADFCQFQOARAVGLSGTFRALSSRLQDLYSIVRRADRGSPV 79  
QY 61 NLKDELLFPSSWEALFSGSEGPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNRRRLTE 120  
DB 80 NLKDEVLSFSDLSFSGSQGLQPGARIFSPDGRDLRHHPAWPKSVWHGSDPSGRRLME 139  
QY 121 SYCETWRTTETGATGQASSLLSGRLLGQSAASCHHAYIVLCIENSFMTA 169  
DB 140 SYCETWRTTETGATGQASSLLSGRLLGQSAASCHHAYIVLCIENSFMTS 188

RESULT 10  
US-09-561-108-13  
Sequence 13, Application US/09561108  
Patent No. 6342221  
GENERAL INFORMATION:  
APPLICANT: Philip E. Thorpe  
TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF  
FILE REFERENCE: 4001.002584  
CURRENT APPLICATION NUMBER: US/09/561,108  
CURRENT FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/131,432  
PRIOR FILING DATE: 1999-04-28  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 13  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-561-108-13

Query Match 87.1%; Score 778; DB 4; Length 191;  
Best Local Similarity 85.8%; Pred. No. 5.6e-87;  
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;  
QY 1 VALNSPLSGMGRGIRGADFCQFQOARAVGLAGTFRALSSRLQDLYSIVRRADRAAIVP 60  
DB 20 VALNTPLSGGMRGIRGADFCQFQOARAVGLSGTFRALSSRLQDLYSIVRRADRGSPV 79  
QY 61 NLKDELLFPSSWEALFSGSEGPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNRRRLTE 120  
DB 80 NLKDEVLSFSDLSFSGSQGLQPGARIFSPDGRDLRHHPAWPKSVWHGSDPSGRRLME 139  
QY 121 SYCETWRTTETGATGQASSLLSGRLLGQSAASCHHAYIVLCIENSFMTA 169  
DB 140 SYCETWRTTETGATGQASSLLSGRLLGQSAASCHHAYIVLCIENSFMTS 188

RESULT 11  
US-09-561-526-13  
Sequence 13, Application US/09561526  
Patent No. 6416758  
GENERAL INFORMATION:  
APPLICANT: Philip E. Thorpe  
TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF  
FILE REFERENCE: 4001.002586  
CURRENT APPLICATION NUMBER: US/09/561,526  
CURRENT FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/131,432  
PRIOR FILING DATE: 1999-04-28  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 13  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-561-499-13

Query Match 87.1%; Score 778; DB 4; Length 191;  
Best Local Similarity 85.8%; Pred. No. 5.6e-87;  
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;  
QY 1 VALNSPLSGMGRGIRGADFCQFQOARAVGLAGTFRALSSRLQDLYSIVRRADRAAIVP 60  
DB 20 VALNTPLSGGMRGIRGADFCQFQOARAVGLSGTFRALSSRLQDLYSIVRRADRGSPV 79  
QY 61 NLKDELLFPSSWEALFSGSEGPKPGARIFSPDGKDLRHPTWPKSVWHGSDPNRRRLTE 120  
DB 80 NLKDEVLSFSDLSFSGSQGLQPGARIFSPDGRDLRHHPAWPKSVWHGSDPSGRRLME 139  
QY 121 SYCETWRTTETGATGQASSLLSGRLLGQSAASCHHAYIVLCIENSFMTA 169  
DB 140 SYCETWRTTETGATGQASSLLSGRLLGQSAASCHHAYIVLCIENSFMTS 188

RESULT 13  
US-09-998-831-13  
Sequence 13, Application US/09998831  
Patent No. 6676941  
GENERAL INFORMATION:  
APPLICANT: Philip E. Thorpe  
TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY  
INHIBITING VEGF  
FILE REFERENCE: 4001.002584  
CURRENT APPLICATION NUMBER: US/09/998,831  
CURRENT FILING DATE: 2001-11-30  
PRIOR APPLICATION NUMBER: 09/561,108

PRIOR FILING DATE: 2000-04-28  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 13  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-998-831-13

Query Match 87.1%; Score 778; DB 4; Length 191;  
Best Local Similarity 85.8%; Pred. No. 5.6e-87;  
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRAPLSSRLQDLYSIVRRADRAAIVP 60  
DB 20 VALNTPLSGGMGRGIRGADFCQFOQARAVGLSGTFRAPLSSRLQDLYSIVRRADRGSV 79  
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPDNGRLTE 120  
DB 80 NLKDEVLSPSWDSLFSGSGQLPGARIFSDGDRVLRHPAWPKSVWHGSDPDGRRLE 139  
QY 121 SYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 169  
DB 140 SYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 188

RESULT 14  
US-08-159-784-2  
Sequence 2, Application US/08159784  
Patent No. 5643783  
GENERAL INFORMATION:  
APPLICANT: Bjorn R. Olsen  
TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,784  
FILING DATE: December 1, 1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: John F. Freeman  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00246/170001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 195  
TYPE: amino acid  
STRANDEDNESS: N/A  
TOPOLOGY: N/A  
US-08-159-784-2

Query Match 86.8%; Score 775; DB 1; Length 195;  
Best Local Similarity 85.2%; Pred. No. 1.3e-86;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRAPLSSRLQDLYSIVRRADRAAIVP 60  
DB 20 VALNTPLSGGMGRGIRGADFCQFOQARAVGLSGTFRAPLSSRLQDLYSIVRRADRGSV 79  
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPDNGRLTE 120  
DB 80 NLKDEVLSPSWDSLFSGSGQLPGARIFSDGDRVLRHPAWPKSVWHGSDPDGRRLE 139  
QY 121 SYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 169  
DB 140 SYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 188

RESULT 15  
US-08-985-526-36  
Sequence 36, Application US/08985526  
Patent No. 6080728  
GENERAL INFORMATION:  
APPLICANT: Mixson, James A  
TITLE OF INVENTION: CARRIER: DNA COMPLEXES CONTAINING DNA  
TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE  
TITLE OF INVENTION: THERAPY  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Connolly, Bove, Lodge, & Hutz  
STREET: 1220 Market Street, P.O. Box 2207  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: U.S.A.  
ZIP: 19899  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/985,526  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/608,845  
FILING DATE: 16-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: McMorrow Jr., Robert G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (302) 658-9141  
TELEFAX: (302) 658-5613  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 185 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-985-526-36

Query Match 82.2%; Score 734; DB 3; Length 185;  
Best Local Similarity 82.4%; Pred. No. 1.3e-81;  
Matches 140; Conservative 14; Mismatches 14; Indels 2; Gaps 2;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRAPLSSRLQDLYSIVRRADRAAIVP 60  
DB 14 VALNTPLSGGMGRGIRGADFCQFOQARAVGLSGTFRAPLSSRLQDLYSIVRRADRGSV 72  
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPDNGRLT 119  
DB 73 QNLKDEVLSPSWDSLFSGSGQLPGARIFSDGDRVLRHPAWPKSVWHGSDPDGRRLE 132  
QY 120 ESYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 169  
DB 133 ESYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 182

PRIOR FILING DATE: 2000-04-28  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 13  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-998-831-13

Query Match 87.1%; Score 778; DB 4; Length 191;  
Best Local Similarity 85.8%; Pred. No. 5.6e-87;  
Matches 145; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRAPLSSRLQDLYSIVRRADRAAIVP 60  
DB 20 VALNTPLSGGMGRGIRGADFCQFOQARAVGLSGTFRAPLSSRLQDLYSIVRRADRGSV 79  
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPDNGRLTE 120  
DB 80 NLKDEVLSPSWDSLFSGSGQLPGARIFSDGDRVLRHPAWPKSVWHGSDPDGRRLE 139  
QY 121 SYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 169  
DB 140 SYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 188

RESULT 14  
US-08-159-784-2  
Sequence 2, Application US/08159784  
Patent No. 5643783  
GENERAL INFORMATION:  
APPLICANT: Bjorn R. Olsen  
TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,784  
FILING DATE: December 1, 1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: John F. Freeman  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00246/170001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 195  
TYPE: amino acid  
STRANDEDNESS: N/A  
TOPOLOGY: N/A  
US-08-159-784-2

Query Match 86.8%; Score 775; DB 1; Length 195;  
Best Local Similarity 85.2%; Pred. No. 1.3e-86;

QY 1 VALNSPLSGMGRGIRGADFCQFOQARAVGLAGTFRAPLSSRLQDLYSIVRRADRAAIVP 60  
DB 20 VALNTPLSGGMGRGIRGADFCQFOQARAVGLSGTFRAPLSSRLQDLYSIVRRADRGSV 79  
QY 61 NLKDELLFPSWEALFSGSEGPKPGARIFSDGKDVLRHPTWPKSVWHGSDPDNGRLTE 120  
DB 80 NLKDEVLSPSWDSLFSGSGQLPGARIFSDGDRVLRHPAWPKSVWHGSDPDGRRLE 139  
QY 121 SYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 169  
DB 140 SYCETWRTTATGATGQASSLLGRLLEQKAAASCHNSYIVLCIENSFMTS 188

us-09-171-607a-1.rai

Sat Mar 13 08:24:05 2004

Search completed: March 13, 2004, 08:19:01  
Job time : 23 secs